uc3m | Universidad Carlos III de Madrid

Bachelor Degree in Biomedical Engineering
2017-2018

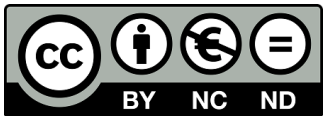
Bachelor Thesis

"STRUCTURAL AND FUNCTIONAL STUDIES ON CAD, THE ANTI-TUMORAL TARGET PROTEIN LEADING DE NOVO BIOSYNTHESIS OF PYRIMIDINES"

María Reverte López

Tutors:
Santiago Ramón Maiques
Sara Guerrero Aspizua

Leganés, 2018



ABSTRACT

CAD is a large multifunctional protein catalyzing the initial three steps in *de novo* biosynthesis of pyrimidine nucleotides in animals. Since pyrimidines are the building blocks of nucleic acids and the precursors of other key macromolecular substances, the up-regulation of CAD's activity is essential for cell growth and proliferation, especially in neoplastic cells. Thus, CAD has been considered an attractive target for the development of anti-tumoral compounds. However, despite the central metabolic role and its therapeutic potential, due to the lack of knowledge about its organization and the structure and function of its different enzymatic domains, no robust inhibitor that could be used as anti-proliferative drug has been designed thus far. In this bachelor thesis, a research study on CAD's DHOase domain was conducted, particularly focused in understanding the role of a flexible loop that appears to play a conserved catalytic role from *E.coli* to humans. To examine the catalytic mechanism of this loop, a cloning approach was designed to generate a human DHOase construct bearing the equivalent flexible loop of the *E.coli* enzyme. This chimeric protein was expressed either in mammalian cells or in bacteria cultures and purified using different chromatographic techniques. Activity assays on both the forward and reverse directions of the reaction were then performed in the chimeric huDHOase to estimate the turnover rate of the mutant. With a negligible enzymatic activity (less than 2% of the wild type) the experiments here presented prove that, despite having a conserved functional role, the flexible loop of *E.coli* and human DHOases are not interchangeable. Overall, the results confirmed the implication of the flexible loop in oligomerization and in the catalytic mechanism of DHOases, highlighting key differences in the functioning of both mammalian and bacterial enzymes that will be exploited in future work for the rational design of specific inhibitors against CAD.

Keywords: Biochemical analysis; Mutagenesis; Recombinant protein; Anti-tumoral target

Glossary

Ca-asp: carbamoyl aspartate
DHO: dihydroorotate
DHOase: dihydroorotase
dNTP: deoxynucleotide triphosphate
ecDHOase: *E.coli* dihydroorotase
FOA: 5-fluoroorotate
FPLC: fast performance liquid chromatography
GFP: green fluorescent protein
HDDP: 4-hydroxyphenylpyruvate dioxygenase
HPLC: high performance liquid chromatography
huDHOase: human dihydroorotase
huDHOase-ELF: human dihydroorotase with *E.coli* loop flexible
IMAC: immobilized metal affinity chromatography
MALS: multiangle light scattering
PALA: *N*-phosphonacetyl-L-aspartate
PCR: polymerase chain reaction
PEI: polyethylenimine
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC: size exclusion chromatography
wt-huDHOase: wild type human dihydroorotase

ACKNOWLEDGMENTS

To Lourdes and Fernando, may this work be the first step to reach the goal I have set since I realised how much you have worked so I could have the opportunities you did not. It might be too early now but at some point I will surely buy you that beautiful and huge villa in Tuscany I always lure you with. With a lake in which Dad only gets sunburned by spending endless days fishing and Mom's only worry is to find a way to take care of thousands of different flowers. If I managed to wake up at six in the morning every day this year is all thanks to your example as the most hard-working parents I have ever known.

To Marina and the crazy human beings belonging to the group chat *Chicuelas*, for listening to my complains about delayed trains, unsuccessful PCRs and colonies-uninhabited plates. For enduring my stressed presence for a whole four years. Four years that, although challenging and tough at times, gift me with moments in which I was able to laugh, cry and live next to you.

To Santiago, Paco and María, the most diligent group at the CBMSO. From the smallest details to the biggest gestures you truly have brought me light during the last year of my college life. Starting with Santiago's daily coffee offers, continuing with Paco's amazing repostery skills and finishing with María's smile always present in her soothing talks and wise advice. Thank you for letting me monopolize part of your bench and overlook at my mistakes, for giving me the opportunity of working with you and gain an immeasurable amount of knowledge from your teachings, your guidance and your words of encouragement. Thank you for letting me discover where my scientific curiosity was all this time and to which field it aims.

[0]and to you reader, I hope you have a nice cup of coffee or tea in your hand ready to enjoy this work and its aclaratory footnotes written with care to ease its reading

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1. Biological Background

1.1.1. Pyrimidine metabolism and CAD

Among the million of species on Earth, the human body dominates as the system that comprises one of the most complex metabolic networks. Through an extensive number of metabolic pathways —still under study—, human cells succeed on carrying out chemical reactions crucial for its growth, proliferation and survivability every second. Among all this highly regulated pathways, the chemical reactions leading the biosynthesis of purine and pyrimidines nucleotides remain as some of the most challenging for research due to its complexity and precedence.

Pyrimidines are heterocyclic compounds with two nitrogen atoms in its ring structure from which several significant organic compounds are derived (Fig. 1.1) [1]. Although these nitrogen bases might appear as basic and revisited, the attention to this type of aromatic compound is such that NASA has lead projects that confirmed the presence of non-terrestrial pyrimidines compounds in meteorites [2]. This discovery, a significant breakthrough for the study on life's origin, did not go unnoticed by the scientific community as pyrimidines constitute the compounds from which, among many others, life's building blocks are derived. Some of the nucleotides precursors that originate from the biosynthesis of pyrimidines, like UDP-sugars and CDP-diacylglycerol diphosphatase, are essential for cell functions like the assembly of cell membranes and the glycosylation of proteins. However, as hinted before, the attractiveness of their study comes from pyrimidines —along with other heterocyclic aromatic organic compounds, purines— being the essential building blocks of crucial molecules like DNA (deoxyribonucleic acid), RNA (ribonucleic acid) or vitamins [3].

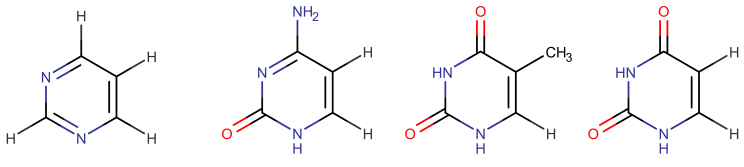


Fig. 1.1. Skeletal formula of the pyrimidine ring and its nitrogen derivatives found in DNA (Cytosine, Thymine) and RNA (Uracil)

Depending on their stage of development and type, cells obtain pyrimidines by two

distinct pathways [4]. On the one hand, by the salvage pathway, quiescent cells[1] obtain pyrimidines in an energy saving approach in which preformed bases and nucleosides, coming from their degradative pathways, are recycled and used as intermediate molecules for the generation of pyrimidines. On the other hand, by the *de novo* biosynthesis pathway, pyrimidines are built from simpler raw molecules like bicarbonate, glutamine and adenosine triphosphate (ATP) through a series of six enzymatic [2] steps highly conserved among all species [5].

In animals, out of the six enzymatic steps for the *de novo* biosynthesis of pyrimidines, the first three reactions are catalyzed by CAD, a 243 kDa multifunctional protein composed by four enzymatic domains that self-assembles into hexamers[3] of ∽1.5 MDa [6] (Fig. 1.2). After Stark *et al.* described how the three last enzymatic domains remain covalently linked and act as a multifunctional protein [7], several studies have described its architecture [6]; the biochemical mechanisms responsible for its behavior and control of its constituents [6]; and, with contradictory results, the localization of CAD in the cell [5].

Although CAD has been reported to be present in the nucleus during some stages of the cell cycle [4], it is widely accepted that the complete four domain construct remains mainly in the cytosol associated with the cytoskeleton network outside of mitochondria [5]. As shown in Fig. 1.3, from bicarbonate, glutamine and two ATP molecules, the glutaminase (GLNase) and carbamoyl phosphate synthase (CPSase-II) domains catalyze the synthesis of carbamoyl phosphate (CP), the first product of the pathway. The next enzymatic domain, aspartate transcarbamoylase (ATCase), channels the labile CP in its carboxyl terminal end inducing its reaction with aspartate and giving carbamoyl aspartate (CA-asp) as product. Lastly, the fourth domain,

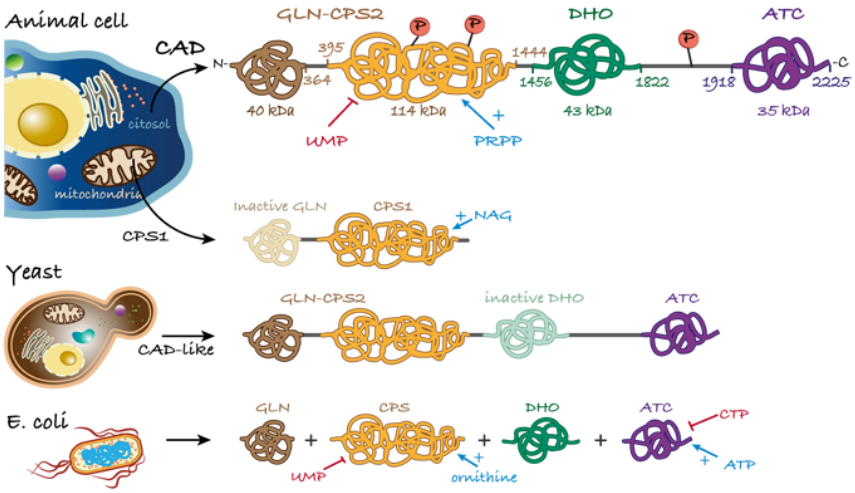


Fig. 1.2. Structural assembly of the proteins involved in the synthesis of pyrimidines by the *de novo* pathway in animal cells, yeast and *E.coli*. Adapted with permission from Grande-García *et al.* [8]

[1] cells halted in a non-dividing state
[2] produced by enzymes, molecular catalysts of living organisms
[3] protein formed by six protein subunits

the zinc metalloenzyme[4] dihydroorotase (DHOase) placed between the CPSase and ATCase domains, catalyzes the reversible cyclization of carbamoyl aspartate to dihydroorotate (DHO), the precursor of the pyrimidine ring [6].

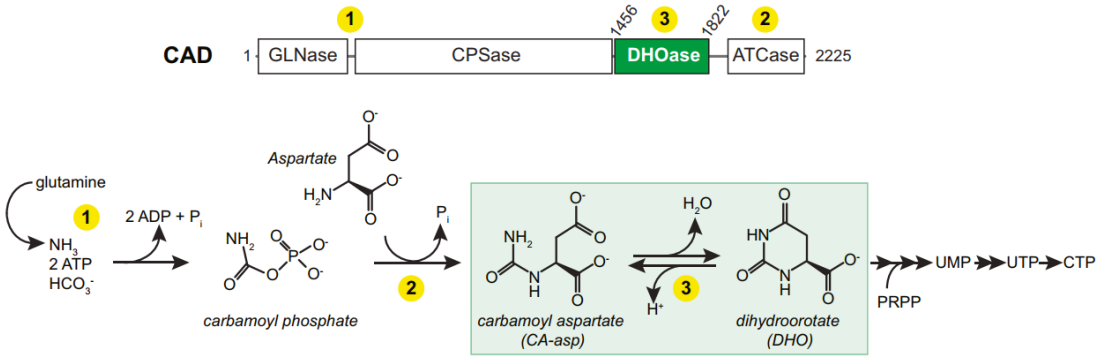


Fig. 1.3. Chemical reactions taking place in the *de novo* pyrimidine biosynthetic pathway. Adapted with permission from Grande-García *et al.* [8]

As CAD carries out the reactions of a high relevance pathway, it must be tightly controlled in order to fit the requirements dictated by cells. By allosteric inhibition[5] through uridine-5'-triphosphate (UTP) and activation by PRPP (phosphoribosyl 5'–pyrophosphate), the CPSase domain regulates the rate at which CAD catalyzes the products of the pathway. In cases in which cells need to grow and proliferate, CAD up-regulates its activity to satisfy the high demand of pyrimidines and its precursors, needed for the synthesis of DNA and other cellular components. However, this up-regulation can also be induced by a disease state like the development of neoplastic cells and tumors, in which the need for rapid division translates into an aberrant use of the *de novo* pathway to synthesize more pyrimidines [5].

Nevertheless, as any metabolic pathway, impairment of its control can lead to maladies. In the case of CAD, as this protein performs such a crucial function to cells, there is only a small amount of documented non-fatal mutations in humans. This mutations, however, are the focus of many studies [3] as they could provide with clues on the interaction between CAD's enzymatic domains and the synergy of their functional roles [9].

1.1.2. DHOase, the lead character

Dihydroorotases (DHOases) are zinc metalloenzymes that take part in the pyrimidine biosynthesis by catalyzing the reversible cyclization of *N*-phosphonacetyl-L-aspartate (CA-

[4]enzymes that contain metal ions providing with some function

[5]negative modulation of a pathway in which a ligand binds to an active site of an enzyme to stop it from functioning

asp) to dihydroorotate (DHO) [10]. The hydrolytic cleavage of DHO is favored at pH 8.0 while the cyclization of CA-asp in the forward reaction is optimal at pH 5.5

By phylogenetic analysis, gene duplication justify the existence of two types of DHOases that have diverged from the same ancestor. Type I DHOases are the most ancient and larger (approx. 45 kDa) and appear covalently linked to other proteins or assembled as enzyme domains of other multifunctional proteins [11]. This is the case of the DHOases found in mammals and Gram-positive bacteria, among others. On the other side, type II DHOases are more recent in the evolutionary scale and present a smaller size (approx. 38 kDa). Type II DHOases appear as monofunctional enzymes that do not interact with the rest of the enzymes of the pyrimidine pathway and present major differences in their sequence with respect to type I [12]. These DHOases are the ones present in eubacteria like *E.coli*, fungi and plants [8].

While members of the same type of DHOases share more than 40% of its sequence identity, when comparing both types, the major changes type II have undergone translates into less than 20% of sequence identity [11]. However, both types belong to the amidohydrolase superfamily, a family of hydrolytic enzymes[6] that share the same folding conformation, the TIM-barrel. This structural framework consists of eight α-helices[7] and eight parallel β-strands[8] that alternate along the peptide backbone forming a solenoid that curves around to close on itself [10].

Depending on the species, DHOases carry different number and types of metal centers; and when this native metals are substituted by Cd^{2+} Co^{2+} and Mn^{2+}, conservation of the catalytic activity is observed. Some studies, by alignment of DHOases sequences between different species, have provided with further identification of the conserved residues that are more likely to be in charge of acting as bridging ligands[9] between the metal cations due to its properties [10]. Nevertheless, the variability around the assembly and number of metals atoms present in the DHOases, is a potential target of study for the design of species-specific inhibitors for this type of enzymes [13].

1.1.3. *E.coli* DHOase

In *E.coli*, the DHOase is a homodimeric[10] enzyme of 76 kDa and 694 amino acids [10]. The crystallization work of Lee *et al.* confirmed the structure of this DHOase. With two identical subunits, each of them folds in a TIM-Barrel conformation (8 parallel β-sheets

[6] enzymes that speed up the hydrolysis of a chemical bond
[7] motif in a protein in which its polypeptide chain coils into a spiral shape
[8] protein motif in which the polypeptide chain shows a streched and sheet-like shape
[9] ligands that connects two or more atoms
[10] protein made by two identical proteins

flanked by 8 α-helices at its outer surface) and presents binuclear zinc centers with both metal ions separated by ~3.5 Å (Fig. 1.4). In both subunits the zinc atoms are coordinated by histidine and aspartate residues. On subunit I, a carbamate functional group (coming from a post translational modification of a lysine), a water molecule, and the DHO product lying close in the active site[11] pocket, bridge the active metals. However, for subunit II, in addition to the carboxylate lysine, one of the carboxylated groups of the CA-asp molecules is also in charge of connecting the zinc atoms [12].

The solved structure for *E.coli* DHOase published by Thoden *et al.*, crystallized in the presence of racemic[12] substrate (D,L-CA-asp), shows the substrate CA-asp bound to subunit II by its active site, while the active site of subunit I is occupied by DHO [10]. Considering this odd and unusual placement of different ligands in each subunit, Lee *et al.* crystallized the enzyme in the presence of only DHO, the product. The crystal unexpectedly showed the same arrangement previously reported, a CA-asp bound to the active site of subunit II while DHO occupied subunit I [13]. To explain this phenomena the authors theorized that, by changes in three pairs of residues close to the dimer interface, each subunit may communicate with the other about its content. However, they also discovered that, even with different ligands in their active sites, the coordination number of the zinc atoms is identical appearing closer to each other only in subunit I [13].

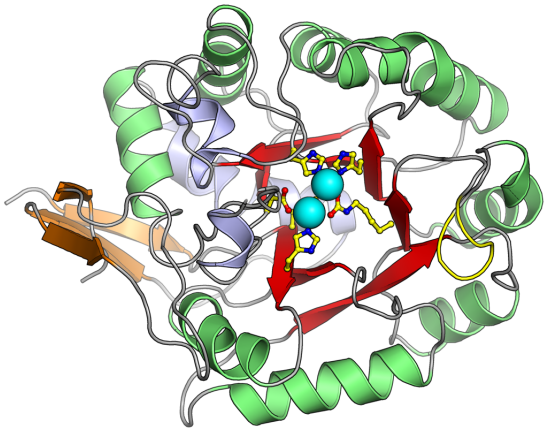


Fig. 1.4. Top view in cartoon representation of ecDHOase showing its characteristic TIM-Barrel conformation and its binuclear zinc center. α-helices are in green, β-sheets in red and the bridging ligands of the two zinc atoms, depicted as blue spheres, are shown in ball-and-stick representation [PDB ID code 1XGE].

[11] region of an enzyme which carries out a chemical reaction when a ligand binds to it
[12] composed by enantiomers of a chiral molecule in equal proportion

1.1.4. The first sighting of a flexible loop

Nonetheless, the most important discovery derived from the ecDHOase crystals analyzed by Lee *et al.* in 2005 consisted on the sighting of a particularly important structure, a flexible loop (annotated as the 4th loop in the 3D model structure) comprised by 11 amino acids (residues from 105 to 115). This loop appears in a different conformation for each subunit of the DHOase from *E.coli*. In the case of subunit II, the loop is a **IN** conformation: the whole structure reaches and curls towards the active site and the two residues at the tip of the loop make hydrogen bonds with the bound substrate CA-asp. Per contra, in subunit I, the loop appears in a **OUT** conformation: no interaction between the loop and the bound product DHO is observed and the loop belongs to the protein surface, facing away from the active site leaving it solvent exposed [12].

This peculiar loop in/loop out mechanism, similar to the movement of a hinge, would explain the catalytic activity of the enzyme as follows [12]:

1. CA-asp starts its cyclization by making direct molecular contacts with the binuclear zinc center of one active site. Due to these metal cations, the C5 of the CA-asp becomes more electrophilic promoting bond formation.

2. The flexible loop comprising residues from 105 to 115 reaches towards the active site inducing a hydrogen bond interaction between Thr109 and the CA-asp. This promotes the optimal and stable confinement of the substrate into the binding site pocket, avoiding its interaction with any water molecule.

3. As now the C4, O4 and O5 of the CA-asp are no longer a carboxylate group but a carbonyl one in DHO, those no longer can act as bridging ligands for the zinc atoms at the active site and causes the DHO molecule to shift its position moving upwards.

4. The movement of DHO away from the active pocket causes pressure on residue Thr109 forcing the loop to adopt its open conformation and leaving the active site cavity filled with water molecules.

Furthermore, this hinge movement happening in both active sites of the protein, could be the source of the communication observed between subunits [12]. The excluding of water molecules during the **IN** conformation and the additional conformational changes happening in the side-chains closer to the loop at the dimer interface, are theorized to cause the asymmetry and communication of both subunits mediating a positive cooperativity with respect to DHO.

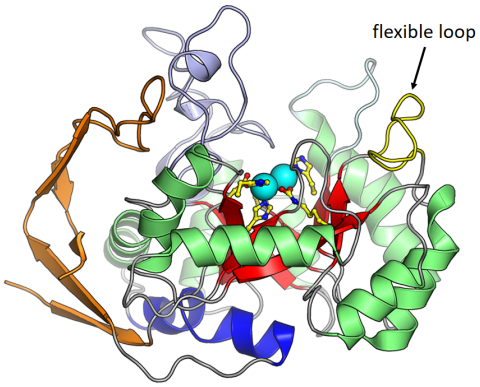


Fig. 1.5. Side view in cartoon representation of ecDHOase showing its flexible loop in yellow, alpha helices are in green, beta sheets in red and the bridging ligands of the two zinc atoms, depicted as blue spheres, in ball-and-stick representation [PDB ID code 1XGE].

1.1.5. huDHOase and its flexible loop

As mentioned before, the human DHOase exists in the cell assembled to other enzymes creating the multifunctional protein CAD. Thanks to the work of Grande-García *et al.*, the structure of the huDHOase has been solved and shows interesting features that set them apart from the other species studied DHOases.

Firstly, the huDHOase, analogous to the other enzymes belonging to the amidohydrolase superfamily, presents a TIM-barrel folding conformation with eight parallel β-strands connected by eight outer α-helices. However, unlike the ecDHOase, the β-strands of the huDHOase are also connected by a smaller domain composed by residues from the N-terminal[13]and C-terminal[13] , the latter crucial for protein solubility [8].

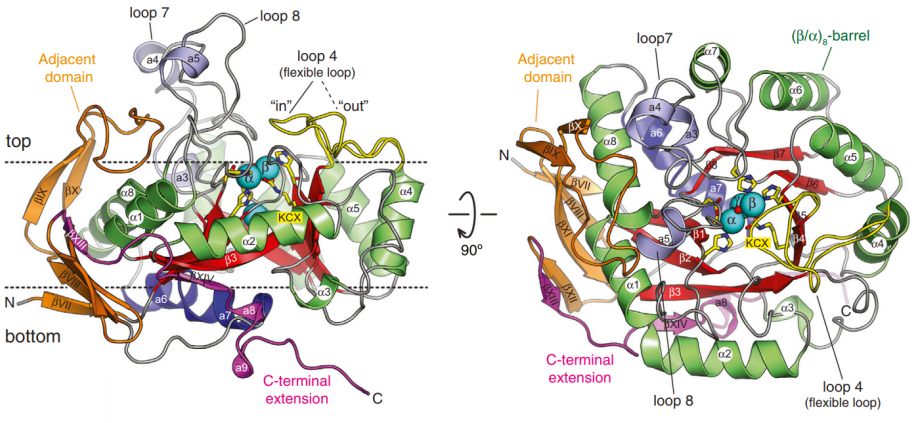


Fig. 1.6. Cartoon representation of huDHOase showing the two loop conformations, its Zn^{2+} in sphere shape and the coordinating elements in ball-and-stick representation [PDB ID code 4C6C]. Adapted with permission from Grande-García *et al.* [8]

[13]termini of a protein, N stands for the initial free amine group of the chain and C for the carboxyl group of the last amino acid of the chain

Regarding its active site, huDHOase (Fig. 1.6) presents a cavity-shaped binding pocket with a major singularity, the presence of three Zn^{2+} ions (Zn-α, Zn-β, Zn-γ). While the Zn-α and Zn-β exhibit their configurations and bridging ligands being identical to the ecDHOase, for the case of Zn-γ its localization has been reported deep into the β-strands of the barrel coordinated by several residue side chains and a water molecule, and connected to the Zn-α by a histidine side chain [8].

In spite of the sequence divergence among the two DHOases, both the human and *E.coli* DHOases maintain a similar architecture and perform its catalytic activity following an analogous approach. By binding of CA-asp to Zn-α and Zn-β through it β-carboxylate group (C4, O4, O5), the two metal cations induce the nucleophilic attack by the N3 atom of the substrate (Fig. 1.3). A proton from this nitrogen is then transferred to O5, which is released as an hydrogen ion, and the bond between C4 and O5 is cleavaged to make the product DHO. For the reverse reaction, as the O4 of the recently generated DHO keeps its bond to Zn-β, this would facilitate the polarization of the C4 and O4 bond making it vulnerable to another nucleophilic attack from the bridging hydroxide ion, leading the catalysis of the substrate in the reverse direction [8].

Once the substrate is bound to the active site, the flexible loop comprised by residues 1561-1569 curls towards the substrate (loop **IN**) and a threonine (Thr1562) interacts with CA-asp stabilizing its conformation and closing the binding pocket avoiding any interaction of the transition state with water molecules. Once the DHO is formed, the loop adopts its **OUT** conformation due to the upwards displacement of the product which forces it to move in a hinge fashion, letting the active site exposed to the rest of the water molecules and disassembling its interactions with the ligand and active site [8].

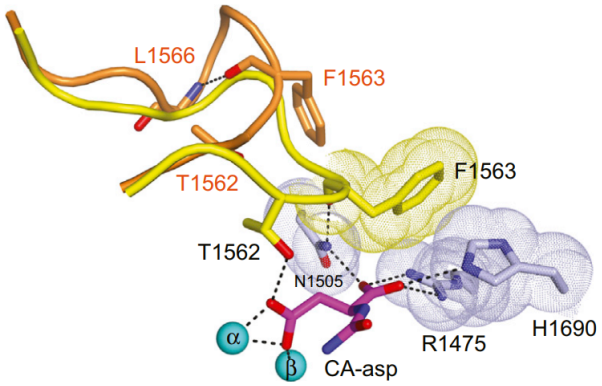


Fig. 1.7. Cartoon representation of huDHOase flexible loop showing the **IN** conformation in yellow and **OUT** conformation in orange. The Zn^{2+} are shown as spheres and the residues responsible of key interactions are labelled. Adapted with permission from Grande-García *et al.* [8]

In the case of huDHOase, studies on the pH dependency of both the forward and reverse catalytic reactions were done following previous experiments performed on *Mus musculus* DHOase [14]. The ionization of the bridging water, and its decrease in pK_a when binding to the metal atoms, translates into two different pK_a for huDHOase: 8.3 in the forward direction and 7.1 for the reverse reaction. These pH variations experiments also demonstrated that, although the Zn-γ does not influence the overall enzyme structure, its binding (dependable on pH conditions) affects the activity of the enzyme [8].

1.2. State of the Art

1.2.1. Studies done by mutating ecDHOase and huDHOase

Extensive study has been done on ecDHOase by mutating isolated residues in order to determine the effect the mutations had on its structure and function. Lee *et al.* performed two single-point mutations that altered the interaction between Thr109 and Thr110 with the substrate CA-asp [11]. These two residues, besides being involved in the binding of the substrate through hydrogen bonds, are also responsible for the stabilization of the enzyme's transition state. When mutations into Ser/Val/Ala/Gly were performed on this two residues, negligible activity was obtained. The rationale behind this result was hypothesized by looking at the crystal structures of the mutants. Despite the absence of a significant change in the overall 3D conformation of the protein, the spatial configuration of the flexible loop and its vicinity appeared altered. The substrate, unable to bind by hydrogen bonds to these two residues due to the mutation, prevented the enzyme from closing the loop and catalyze the product. One one side, Thr109 was hypothesized to contribute in enhancing substrate affinity, while Thre110 could be responsible of subtrate binding. Furthermore, by deletion of the whole flexible loop (residues from 107 to 116) ecDHOase mutants were assayed and no activity was observed. Therefore, although non-directly, the flexible loop was proposed as one of the key elements of the enzyme that would mediate catalysis by enhancing the affinity for the substrate and the stability of the transition state [11].

Aside from the results, in their methodology Lee *et al.* described the use of the QuikChange Site-directed Mutagenesis[14] kit from Stratagene for the mutations of the residues using a pBS+ plasmid[15] with ecDHOase as template. To purify the protein, after being produced by expression in *E.coli* strain X7014a, the lysate of the cells was subjected to thermal denaturation and MonoQ ion exchange columns and Microsep centrifugal devices were used for the purification and concentration of the protein. For the activity

[14] procedure in which the genetic material of an organism is changed

[15] circular DNA strand that is present in bacteria cytoplasm and replicates separately from its chromosomes, used extensively for the manipulation of genes in the laboratory

assays, spectophotometric measures at 230 nm following the formation of product DHO were performed in 96-well plates, and the reaction was triggered by the addition of the mutant enzymes by a Mosquito nanoliter liquid handling robot [11].

Regarding huDHOase, several mutations reported by Grande-García *et al.* aided in a preliminary hypothesis on the structures and functional groups crucial for the protein dimerization, the binding of the zinc atoms and its catalytic function. Moreover, residues of the human flexible loop were mutated following the results obtained from Lee *et al.* [11] to test their relative activity. The mutations of T1562 and F1563 (residues equivalent to Thr109 and Thr110 in EcDHOase) into Ala showed once again the inactivation of the enzyme (Fig. 1.8). This could support the hypothesis describing the flexible loop as an essential component for the catalysis. However, no further mutations that could possible unravel new interactions between the flexible loop and the ligand in the binding pocket were performed [8].

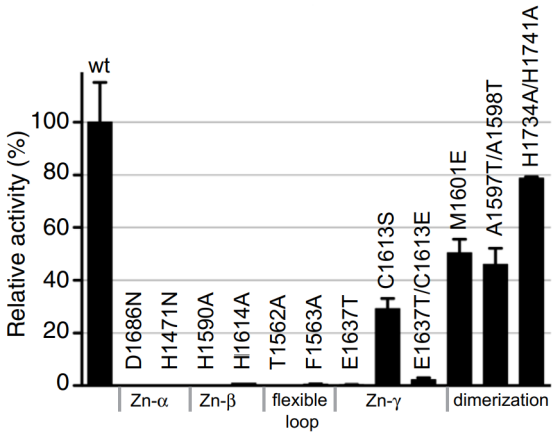


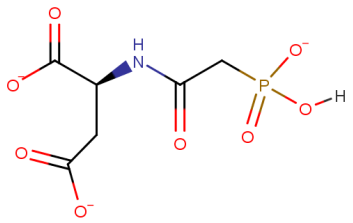
Fig. 1.8. Relative activity of the mutant huDHOases performed by Grande García *et al.* showing almost no activity for the mutations correspoding to the flexible loop. Adapted with permission from Grande-García *et al.* [8]

The methodology followed by Grande-García *et al.* also included the use of the QuikChange Site-directed Mutagenesis for the point mutations of the huDHOase. The plasmid taken as template, in this case, was pOPIN-M and the cells used to express the plasmid, HEK293S-GnTI$^-$. To purify the protein obtained, Ni^{2+} HisTrap Chelatin FF and HiTrap SP HP columns (GE Healthcare) were used before a last step consisting of size-exclusion chromatography with a Superdex 200 10/300 column (GE Healthcare). Once the protein was purified, they employed a 10 kDa cut off membrane of the Amicon Ultra system. In the case of the activity assays, a colorimetric approach was adopted in a 96-well plate in which the formation for product DHO was quantitatively measured following its absorbance at 230 nm [15].

1.2.2. Studies on CAD's possible inhibitors

The search for a molecule able to bind to any of CAD's enzymatic domains in order to halt the *de novo* pathway of pyrimidine biosynthesis started with Collins and Stark in 1971. These two scientists, after an extensive study on the ATCase structure, synthesized PALA (*N*-phosphonacetyl-L-aspartate, seen in Fig. 1.9) a synthetically designed molecule with high potential as a stable inhibitor of ATCase [7].

PALA is an analogue molecule of the transition state, a tetravalent hydrophobic molecule that is susceptible to cellular uptake by fluid phase endocytosis[16]. It was designed with structural features from both substrates of the ATCase in order to robustly bind in the active site and leave the enzyme halted in a conformation similar to the one established during the transition state [16]. Studies on its efficiency report a binding strength 1000 times higher than the binding strength of substrate carbamoyl phosphate and its ability to block the proliferation of a homogeneous solution of cells *in vitro* by its administration on a high dose [17].



PALA

Fig. 1.9. Skeletal formula of *N*-phosphonacetyl-L-aspartate (PALA), an ATCase inhibitor [7].

On clinical trials, the most important evaluation for the anti-tumoral drug PALA showed inhibition of 16-72% of the ATCase activity in tumor and normal tissue, thus inhibiting the *de novo* pathway of pyrimidine biosynthesis [17]. What is more, upon restoration of the pathway by removal of PALA, no pyrimidine synthesis was observed indicating an unexplained prolonged retention of its effect by the cells [18].

On the other side, all *in vivo* assays in tumor tissue showed also how the pyrimidine metabolic behaviour of cells vastly depends on the dose administered and, even when the maximum tolerated dose was administered, PALA is not able to decrease the pool production of UTP and CTP to the extent of completely inhibiting the *de novo* synthesis. However, one of the most promising observations coming from this studies, is the ability of PALA to selectively affect only the tumor tissue while sparing the rest of healthy cells. Consequently, as malignant cells have limited survival ability under such adverse

[16]process in which cells internalize external matter bound to vesicles into its cytoplasm by energy expenditure

condition, PALA would aid in the delay of the tumor growth and mean survival time if combined with other anti-tumoral drugs as its use as a single therapeutic strategy does not show promising results [17].

Regarding the DHOase, a particularly interesting study done in 2007 by Lee *et al.* [19] presented their discoveries on ecDHOase as a precedent of possible research on the DHOase from the Plasmodium species, parasites involved in malaria disease. These four types of Plasmodium responsible for the disease present 6 different enzymes to synthesize pyrimidines through the *de novo* pathway and are unable to carry out the salvage pathway, limiting its resources of pyrimidines from the *de novo* pathway. This revelation could open a new line of research in the future [19]. If the DHOase from this Plasmodium has a custom-built structure or a different catalysis action, then specific inhibitors against this DHOase (sparing the huDHOase) could be synthesized and be tested as a possible treatment against malaria.

In the case of ecDHOase and huDHOase, 5-fluoroorotate (FOA) has been used as a product-like inhibitor in the crystallization work of Lee *et al.* and Grande-García *et al.* to study the interactions of this molecule with the active site. Precluding the enzyme from hydrolyzing this inhibitor, the conformation of the ecDHOase and huDHOase flexible loop stayed with and **OUT** conformation, leaving FOA bound in the active site. However, the use of this product-mimic in both studies is only limited to enable the formation a stable tetragonal crystal structure for the crystallization of the proteins [8] .

To test the power of FOA as inhibitor of Plasmodium DHOase, the study by Gomez and Rathod [20] presented data on the effectiveness and toxicity of this inhibitor on malaria infected mice. With a dose dependent administration of FOA, infected mice showed no sign of parasitemia in the case of the higher dose (5 mg/kg). Despite this good result, several mice died in the process, not from the disease, but from the toxicity of this dose. To counteract this, uridine was administered to the mice in order for the huDHOase to synthesize pyrimidines by the salvage pathway. These mice, with both a high dose of FOA and an additional administration of uridine, were cured of the malaria and did not present any adverse effect related to the toxicity of the inhibitor [20].

The last inhibitor for huDHOAse proposed to date is an ester variation of HDDP (Fig. 1.10), 4-hydroxyphenylpyruvate dioxygenase, a derivative of of pyrimidine-4-carboxylic acid created by Christopherson *et al.* [21]. Although no clear description on the transport mechanism has been described yet, when tested in human leukemia samples, promising results were obtained. No cytotoxicity was observed on both mouse and human leukemia grown in culture. With no major perturbation in the intermediates of the nucleotide metabolism, at a concentration of 250 μM it induces the death of human CCRF-CEM

leukemia cells [22]. However, as promising as this might sound, the results presented by the authors were not obtained as a consequence on the halting of the *de novo* biosynthesis pathway. Inhibition of the DHOase was not observed and the administered HDDP-variant affected the cells by inhibiting their cycle progression from late G_2 or M phase to G_1 leading the inhibition of their growth [22]. Besides HDDP and its variants, no other feasible inhibitor against huDHOase has been proposed for experimentation in vivo which, in combination with PALA, could allow a superior performance in the blockage of the *de novo* pathway.

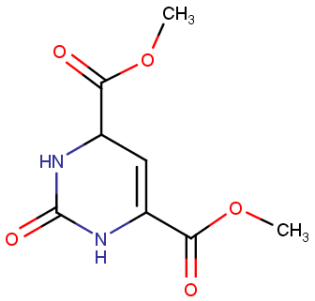


Fig. 1.10. Skeletal formula of HDDP ester variant, a possible inhibitor for huDHOase [22]

1.3. Regulatory Framework

The experimental work performed for the accomplishment of this project was all done under the regulations concerning the chemical and biological practice European and Spanish Royal Legislative Decrees dictate.

Royal Decree 665/1997, eventually modified in R.D 1124/2000, provides the directive to follow during the manipulation, storage and transportation of the chemical reagents used in this project for the mutagenesis and chromatography procedures. Moreover, according to the Centro de Biología Molecular Severo Ochoa (CBMSO)'s requirements, β-mercaptoethanol, acrylamide and other special chemical compounds were treated with particular care to comply with the articles drawn up attending the prevention of risks derived from the use of chemicals, as well as the restrictions applied on their management.

Furthermore, the Royal Decree 833/1988, modified later in R.D 952/1997, establishes the regulation that must be taken into account when managing heavy metals and their treatment to ensure the protection of the of human health, the protection of the environment and the preservation of the natural resources. As in this project nickel was used, its storage and processing was approached by compliance of both the Royal Decree and the CBMSO's chemical residues policy.

As cell culture related practices were performed in a Biosafety Level 2 (BSL-2) laboratory, this working environment is regulated by some postulates in the frame of the Royal Decree 664/1997 for the protection of workers against the biological risks related to the exposure to biological agents. Firstly, for the grant access of the BSL-2, a mandatory training curse conducted by the CBMSO for the correct use of its facilities was undertaken. Good practice for the minimization of the liberation of the biological agents, along with its storage and evacuation and the personal hygiene when exposed, were followed to comply with both the CBMSO policy and consequently, with the Spanish legislation. However, despite mutating the sequence of a protein by cloning techniques and producing the protein in cells, no special additional regulations are involved for this part of the project as the article 1 of the Royal Decree 178/2004 for the handling of mutated biological organisms determines as exception this type of mutagenesis *in vivo*.

Although Biotium (the manufacturer of the nucleic acid gel staining GelRed) conducted several tests with three independent testing services to determine the non-mutagenic and non-cytotoxic character of the dye, as the working conditions in the CBMSO involved the use of equipment handled by other groups that did not use GelRed but Ethidium Bromide (EtBr) or other carcinogenic dyes, the procedures performed in this project followed the requirements for safety and health protection against carcinogenic agents prescribed by the Directive R.D 665/1997. This Royal Legislative Decree, modified later in R.D 1124/2000, formulates the regulations laboratory personnel must follow when handling or being exposed to carcinogenic agents. In the frame of the work done in this project, several approaches for the safety of both the student and the rest of CBMSO personnel were adopted in accordance to articles 4, 5 and 6. By using GelRed instead of EtBr the project abided by article 4 that seeks for the use of less dangerous substitutes, article 5 (Prevention and Reduction of Exposure) was guaranteed with the special treatment of the gels after its use in a separate container for its eventual processing, and finally for article 6 personal hygiene and action lines for the individual protection were followed with the use of gloves and coat.

Finally, the handling of the non-chemical, non-biological and non-carcinogenic residues derived by the laboratory work, was followed by the Law 5/2003 of the Autonomous Community of Madrid and the measures for the occupational health and safety of the laboratory personnel were applied in accordance with Law 54/2003.

2. MOTIVATIONS AND OBJECTIVES

As mentioned in the introductory section of this work, impairment of the complex human metabolic network causes numerous diseases. But, from Alzheimer [23] to Huntington's disease [24], enzyme inhibitors that handicap abnormal enzyme's catalytic activity have proven to be a feasible alternative in increasing the efficacy of current therapies or becoming the only therapeutic regime in the future.

Regarding cancer, the heterogeneity of this disease along with current studies envisaging it as one of the lead causes of death in the next decades [25], are two powerful incentives in the motivation for this project, as both encourage the search for alternative treatments to the traditional cytotoxic procedures like chemotherapy and radiotherapy. Therefore, by using the current progress molecular biology has exhibited in the last years, stratified medicine and personalized strategies for patients are the two interests engaging in the battle against the large amount of different types of cancer.

Furthermore, with the modest research done on human CAD and its domains, every study on its structure or activity is an additional step towards the complete understanding of human's pyrimidine synthesis pathway. And, once the key elements of the pathway are determined, the discovery of new approaches to halt it will follow.

With that being stated, three main motivations leaded the study of huCAD's DHOase domain:

- The need to transfer to academia data about huDHOase's elements in order to find its weaknesses and vulnerabilities as enzyme.
- The demand from pharmaceutical companies on better inhibitors for the huDHOase domain.
- The search for an alternative/additional anti-tumoral treatment to fight cancer.

Motivated by those three *catalysts*, the objectives set in this project for the accomplishment of at least the first one are:

1. To design and choose the best experimental strategy to enable the cumbersome mutation of the flexible loop of huDHOase into the flexible loop from ecDHOase in a plasmid.
2. Express the plasmid in mammalian cells and produce the protein.

3. Purify the protein by selecting the proper chromatography techniques.
4. Test experimentally, by activity assays, the catalytic activity of the new mutant: huDHOase-ELF.
5. Draw conclusions on the key elements of huDHOase sequence responsible on its catalytic function.
6. Propose techniques and procedures needed to solve the possible problems encountered during the project and present the future work required to continue the study on huDHOase.

Ultimately, as shown in Fig. 2.1, the main goals of the project were the successful completion of 4 experimental objectives that could then generate the data needed to analyze the mutations' repercussions and their significance.

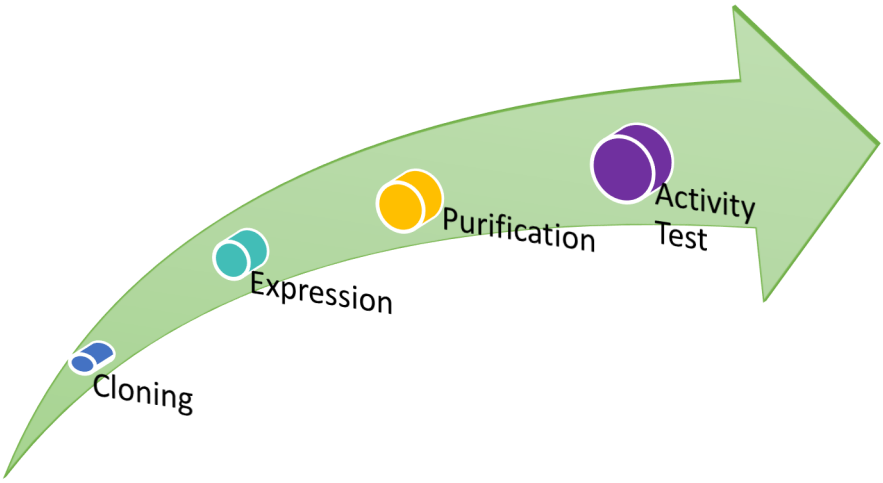


Fig. 2.1. Diagram of the experimental goals set for the project.

3. MATERIALS AND METHODS

In this chapter, the main procedures performed to reach the goal proposed are presented in separate sections in which innovative approaches, and common experimental protocols, are explained. For the sake of clarity, a short description of the workflow is as follows:

I. Linearization, by digestion with restriction enzymes, of the circular plasmid pOPIN-M-huDHOase that contains the huDHOase (needed as template for the cloning procedure) and other important features (6xHis tag, MPB protein and PS site) for the future purification of the mutant portein.

II. Once a set of primers[17] was designed and prepared, as the sequence selected to be mutated was remarkably long, the cloning step was divided into 4 rounds of separate and consecutive mutagenesis protocols with a pair of mutagenic primers for each round. In the first and last rounds, one mutagenic strategy was performed (The Three PCRs Method) while in the second and third rounds a completely different approach was taken (QuikChange Site-directed Mutagenesis).

III. After each round of mutagenesis, the plasmids containing the gradually mutated huDHOase sequence were verified by NaeI digestion and sequencing so cloning could continue progressively with the confirmation that the mutations were present.

IV. To produce the mutant huDHOase-ELF in a high throughput manner, transfection of mammalian cells with the final mutated plasmid was performed with polyethylenimine[18] (PEI).

V. To isolate the protein of interest from the rest of proteins produced by the cells, purification was accomplished using IMAC[19] (Immobilized Metal Affinity Chromatography) and SEC[20] (Size Exclusion Chromatography).

VI. Finally, to assess the functionality of the purified mutated protein, activity assays were carried out on both the wild type huDHOase and chimeric huDHOase-ELF to examine the effects the mutations had on its catalytic activity.

[17] short DNA strand that provides with a starting point to the DNA polymerase for the synthesis of DNA

[18] non-viral vector used in the transfection of nucleic acids into cells

[19] chromatography technique that separates proteins by their different affinity to metal ions

[20] chromatography technique in which molecules are sorted by size due to their different efficiency penetrating the pores of the stationary phase inside a column

3.1. Cloning of huDHOase

3.1.1. huDHOase-pOPIN-M plasmid digestion

In this project, a special construction of the plasmid pOPIN-M (Oxford Protein Production Facility) was used. This plasmid consisted of the wild type human DHOase sequence, a PreScission Protease cleavage site, an ampicillin resistance gene and the sequences coding for the maltose binding protein (MBP) and a tail of histidines (6xHis tag) (Fig. 3.1). Any of these features, essential components needed to perform the cloning, transformation and purification protocols.

To digest the circular plasmid pOPIN-M-huDHOase for further use in the In-Fusion Reaction, restriction enzymes were used to cut the double stranded genetic material into a linear strand. For that, a reaction mix with 1662.5 ng/μl plasmid pOPIN-M-huDHOase, NEBuffer 10X (New England Biolabs), BSA 100X (New England Biolabs) and MilliQ water was prepared and, separating this mix in different tubes, one type of restriction enzyme (either HindIII or KpnI from New Englad Biolabs) was added in a 1:0.2 relation (μg of Plasmid:Units of enzyme). After incubation at 37ºC for 3 hours to assure the correct linearization of the vector, the aliquots were kept frozen at -20ºC until their further use.

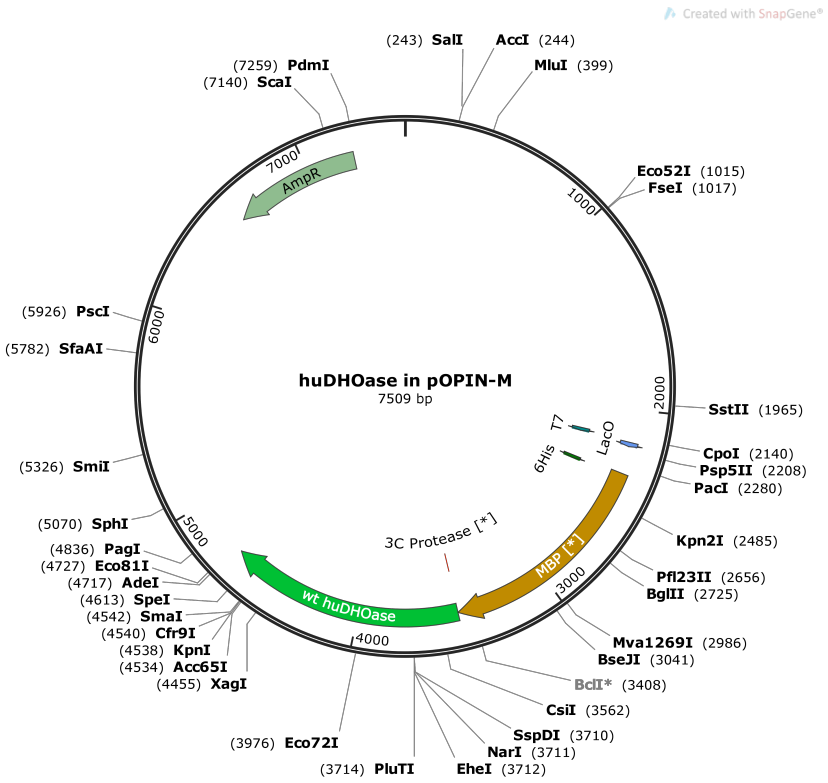


Fig. 3.1. Map of wild type huDHOase in pOPIN-M.

3.1.2. Primer design and preparation

To start the mutation process, the nucleotides conforming the flexible loop from human DHOase were first analyzed (their implication in dimer formation, localization, type, etc.) and several of them were chosen to be mutated in the cloning procedure. In this case, the sequence mutated was 11 amino acids in length and belongs to the flexible loop domain of each subunit of the huDHOase homodimer.

The total amount of 33 nucleotides were mutated in what will be addressed as rounds. In each round, a different pair of primers (one primer for the direct elongation *d* and another for the reverse *r*) were used to mutate and insert the desired bases. The aim behind their sequence design containing the sought mutations was for the DNA polymerase to elongate them in a polymerase chain reaction (PCR)[21] ultimately creating copies of mutated DNA chains.

The design of those primers was done taking into account both the target sequence (human) and the one pursued (*E.coli*):

Human DHOase flexible loop	L N E – T F S E L – R L D
E.coli DHOase flexible loop	P A N A T T N S S H G V T

As indicated in Table 3.1, for the first round, the first three amino acids (L N E) were mutated into the first three of the *E.coli* sequence (P A N). In the second round, the previous DNA sequence was used to continue the transformation into *E.coli*'s loop with an insertion of Alanine and the mutation of the next two amino acids into Threonine-Threonine. For the third round, by taking the sequence coming from round 2 as template, the three consecutive amino acids of the huDHOase were mutated into Asn-Ser-Ser. Finally, in the fourth round, the previous mutated DNA was used as template again and an insertion of Histidine and the mutation of the next three amino acids was performed (H G V T).

	Amino acids targeted in huDHOase for mutation	Amino acids of E. cDHOase inserted in the primer	Primer Pair
Mutagenesis Round #1	CTC GCC AAC	CCC GCC AAC	MR-mut1-d
	L N E	P A N	MR-mut1-r
Mutagenesis Round #2	ACC TTC	GCC ACC ACC	MR-mut2-d
	T F	A T T	MR-mut2-r
Mutagenesis Round #3	TCT GAG CTG	AAC AGC AGC	MR-mut3-d
	S E L	N S S	MR-mut3-r
Mutagenesis Round #4	CGG CTG GAC	CAC GGC GTG ACC	MR-mut4-d
	R L D	H G V T	MR-mut4-r

TABLE 3.1. TARGET SEQUENCES AND MUTATING BASES COMPRISING THE CENTRAL PART OF EACH PRIMER PAIR

[21] laboratory technique used to generate millions of copies of a template DNA strand

Mutagenesis in consecutive rounds was possible due to the architecture given to the primers, and their peculiar attachment to the target sequence in the annealing[22] step of a polymerase chain reaction. As seen in Table 3.2, primers that will be used to mutate huDHOase flexible loop have two distinct constituents. On one side, two extremities, represented in green color, were incorporated in the primer as they are complementary to the bases of the target huDHOase and will anneal to this sequence when subjected to the right temperature in a PCR. On the other side, represented in red color, the mutating bases that will give rise to the ecDHOase loop are integrated in the middle and, although they won't anneal to the target sequence, the stretch of these bases will remain in place thanks to its bound extremities. In this way, when a PCR reaction takes place, the DNA polymerase[23] will start attaching nucleotides to these primers until the rest of the template is completely copied, creating a mutated chain due to the presence of the ecDHOase bases (coming from the middle of the primers) in the final sequence.

Table 3.2 also includes the specifications of DHOd and DHOr3. For these two primers, blue bases correspond to complementary nucleotides of the plasmid pOPIN-M while green colored bases represent the complementary to the huDHOase initial and final sequences.

Furthermore, to avoid the formation of secondary structures in the primers that would prevent them from properly annealing into the template sequence during PCRs, the melting temperature was taken into account and only the primers with the lowest melting temperate were chosen [26].

Primer name	Length	Mw (g/mol)	Tm (ºC)	Sequence of Primers 5'-3'
MR-mut1-d	39	11880	70.2	GGG CTG AAG CTT TAC CCC GCC AAC ACC TTC TCT GAG CTG
MR-mut1-r	39	12098	70.2	CAG CTC AGA GAA GGT GTT GGC GGG GTA AAG CTT CAG CCC
MR-mut2-d	39	11786	73.7	CTT TAC CCC GCC AAC GCC ACC ACC TCT GAG CTG CGG CTG
MR-mut2-r	39	12195	73.7	CAG CCG CAG CTC AGA GGT GGT GGC GTT GGC GGG GTA AAG
MR-mut3-d	39	11927	75.4	GCC AAC GCC ACC ACC AAC AGC AGC CGG CTG GAC AGC GTG
MR-mut3-r	39	12055	75.4	CAC GCT GTC CAG CCG GCT GCT GTT GGT GGT GGC GTT GGC
MR-mut4-d	43	13209	74.2	ACC ACC AAT AGC AGC CAC GGC GTG ACC AGC GTG GTC CAG TGG
MR-mut4-r	43	13243	74.2	CCA CTG GAC CAC GCT GGT CAC GCC GTG GCT GCT ATT GGT GGT
DHOd	41	12576.2	68.6	AAG TTC TGT TTC AGG GCC CG A TGA CCT CCC AAA AGC TTG TG
DHOr3	37	11477.5	64.3	ATG GTC TAG AAA GCT TTA ATC AGG AAG CCC TGG GAT G

TABLE 3.2. FEATURES OF THE PRIMERS DESIGNED

After purchasing the primers from Sigma-Aldrich, they were all prepared in a 20 μM concentration by resuspending them in TE 1X pH 8 (10mM Tris-HCl, 1mM EDTA) and

[22] pairing attachment of complementary bases on two single stranded DNA sequences through hydrogen bonds to create a double stranded sequence

[23] enzyme that synthesize DNA by adding nucleotides to a pre-existing short nucleotide strand already annealed to the template

incubating them 10 minutes for hydration.

3.1.3. The Three PCRs Method

For the first and fourth round of mutagenesis the Three PCRs Method was used. The name given to this procedure comes from the fact that three polymerase chain reactions (PCRs) must be performed to successfully complete one round of mutagenesis. A PCR is a procedure used in molecular biology in which a strand of DNA can be *amplified* to obtain million of copies, due to the ability of DNA polymerases to add nucleotides to a preexisting small piece of DNA called primer which is complementary to the beginning of the template region to be amplified. In the case of the experimental technique here presented, PCRs were performed but not only for the amplification of the strand of DNA. Exploiting PCRs basic principle, another different purpose was sought: to first mutate and then amplify a DNA sequence.

In order to explain the method in an precise way, the first round of mutagenesis will be used as example to determine which reagents and components are needed in the process and which outcomes are expected from each reaction. Nevertheless, to perform the fourth round of mutagenesis, the same protocol was followed and the only change applied was the pair of primers used to mutate the sequence.

As seen in Fig. 3.2, the first two steps applied in The Three PCRs Method for the first round of mutagenesis consist of:

- PCR1: a PCR performed with a high fidelity DNA polymerase that elongated two pair of primers (MR-mut1-r and DHOd) by using as template the huDHOase in pOPIN-M. With this PCR1, a small piece of the whole huDHOase sequence is expected to be amplified, but with the singularity that this PCR1 product contains at its 3' end the particular mutations given by the primer MR-mut1-r that, despite containing mutations in the middle of its sequence, annealed properly to the template.
- PCR2: a PCR performed with a high fidelity DNA polymerase that elongated a different pair of primers (MR-mut1-d and DHOr3) by using as template the huDHOase in pOPIN-M. In this case, a bigger and different sequence corresponding to the end of the huDHOase template is amplified with the mutations the primer MR-mut1-d introduces when extension of its sequence happens.

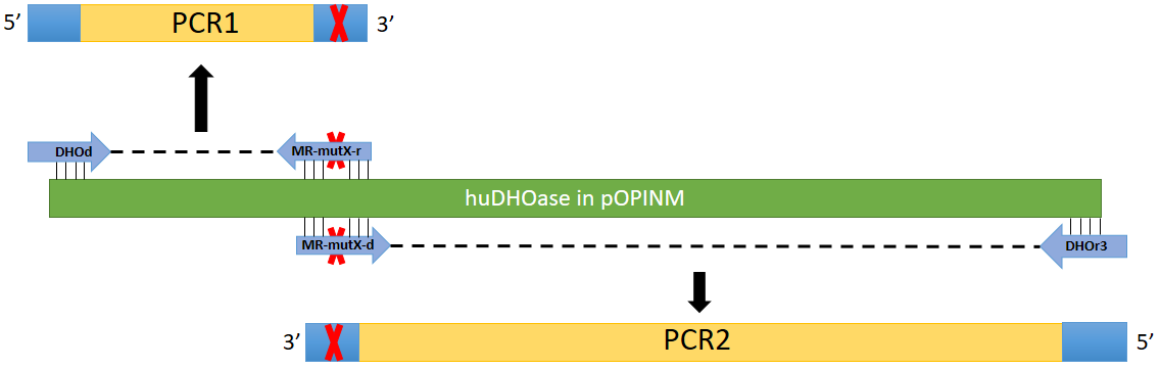


Fig. 3.2. Diagram representing the process behind the synthesis of both PCR1 and PCR2 products from huDHOase in pOPIN-M (depicted in green). Arrows represent the primers that direct the elongation and contain the mutations desired (red crosses) to make two mutated DNA fragments different in size represented in yellow, PCR1 and PCR2.

The reaction mix of 100 μl prepared for PCR1 consisted of: 20 μM of primer DHOd (Sigma), 20 μM of primer MR-mut1-r (Sigma), 10 mM deoxynucleotide triphosphate (dNTPs; ThermoFisher), Buffer Phusion 5X (New England Biolabs), 2u Phusion polymerase (New England Biolabs) and 100 ng of pOPIN-M-huDHOase. In the case of PCR2, the 100 μl reaction mix consisted of: 20 μM of primer DHOr3 (Sigma), 20 μM of primer MR-mut1-d (Sigma), 10 mM dNTPs (ThermoFisher), Buffer Phusion 5X (New England Biolabs), 2u Phusion polymerase (New England Biolabs) and 100 ng of pOPINM-M-huDHOase.

For both PCR1 and PCR2, the conditions used in the thermal cycler[24] (SimpliAmp Thermal Cycler, ThermoFisher) were as presented in Fig. 3.3 and their reaction mix separated into two tubes of 50 μl to allow for a more uniform change of temperature in the solutions.

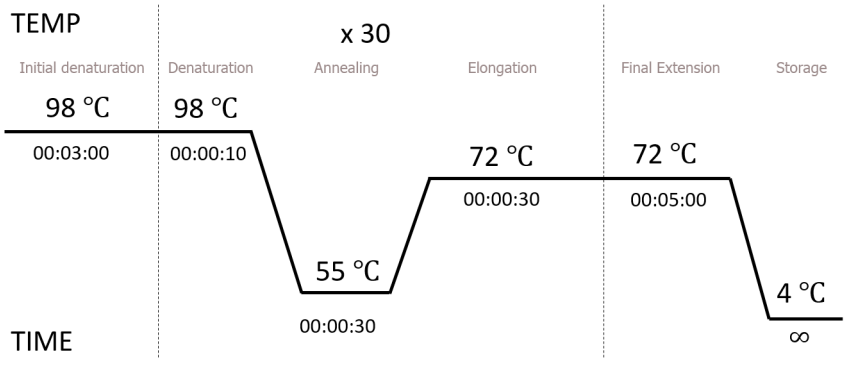


Fig. 3.3. Thermal Cycler conditions used to obtain PCR1 and PCR2.

Once the 30 cycles of the PCR were finished, the two 100 μl reactions were loaded in

[24]laboratory instrument that expose a PCR reaction to different temperatures during programmed periods of time and cycles inducing the reagents of the reaction to amplify the DNA

a 1.2% agarose (CONDA) gel [27], stained with GelRed (Biotium) and observed under UV light with a Slite 140S transilluminator (Pacific Image Electronics) for agarose gel electrophoresis[25]. Subsequently, the bands corresponding to PCR1 and PCR2 that ran up to the expected sizes, were cut and purified with a QUIAGEN Gel Extraction Kit.

The last step to complete the Three PCR Method presented in Fig. 3.4, is a final PCR addressed as PCR3:

- PCR3: PCR performed with a high fidelity DNA polymerase using primers DHOd and DHOr3 to elongate products PCR1 and PCR2 that now act as templates of each other. These two fragments will anneal in the PCR reaction as they share a complementary (and mutated) piece of sequence on their 3' ends. When annealing of both the products and the primers takes place, elongation with the DNA polymerase occurs and the final product addressed as PCR3 results in the complete and non-fragmented huDHOase sequence mutated with the bases sought in the middle.

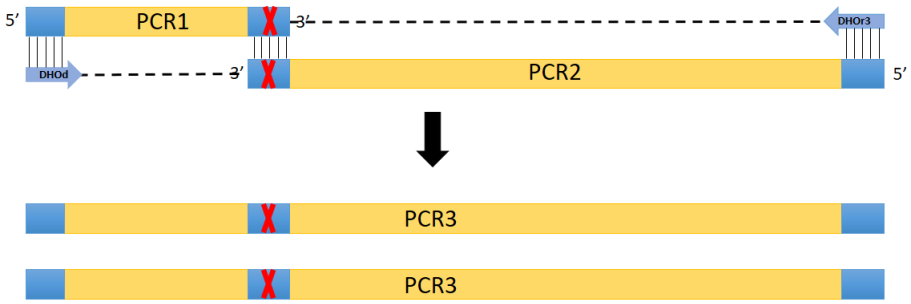


Fig. 3.4. Diagram showing how, from a pair of primers annealed at the ends of both PCR1 and PCR2 products, the complete huDHOase sequence is synthesized with the mutations sought (red crosses).

An important detail to take into account is that, as the PCR1 product is 2.5 shorter in length than the product from PCR2, the reaction must contain double the amount of PCR1 product in comparison with the concentration of PCR2 product added. Therefore, the 100 μl reaction mix consisted of: 20μM of primer DHOd (Sigma), 20 μM of primer DHOr3 (Sigma), 200 ng PCR1 product, 100 ng PCR2 product, 10 mM dNTPs (ThermoFisher), Buffer Phusion 5X (New England Biolabs), 2u Phusion polymerase (New England Biolabs) and 5% DMSO (ThermoFisher).

As the PCR3 requires only the use of primers DHOd and DHOr3, as seen in Fig. 3.5 the temperature of annealing in the Thermal Cycler was set to 60ºC as this primers present the most efficient annealing at this temperature. However, as the same DNA polymerase

[25]technique used for the separation of DNA fragments according to their size by their differential migration in an agarose matrix

was used, the denaturation and elongation temperatures were kept the same as the ones used for PCR1 and PCR2. The last change implemented was the numbers of cycles (35) in order to increase the amount of amplified product.

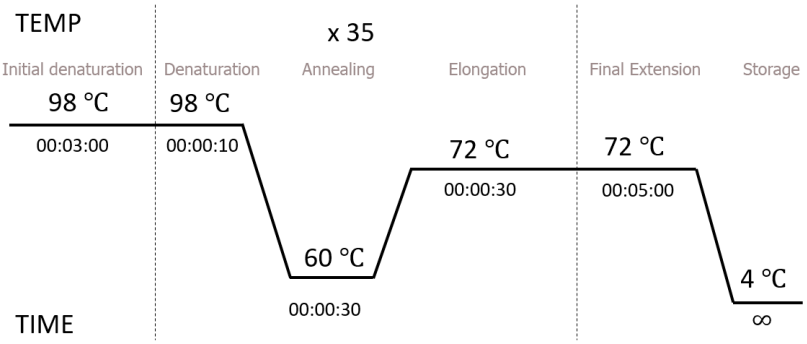


Fig. 3.5. Thermal Cycler conditions used to obtain PCR3.

Once the 35 cycles of the PCR were finished, the 100 µl reaction was loaded in a 0.8% agarose (CONDA) gel and stained with GelRed (Biotium) [27]. To observe the stained band under UV light, a Slite 140S transilluminator (Pacific Image Electronics) was used and subsequently, the band that ran up to the expected size for PCR3, was cut and purified with a QUIAGEN Gel Extraction Kit.

3.1.4. Temperature gradient tests

Each primer used in this project (DHOd, DHOr3, MR-mutX-d, MR-mutX-r) required for a specific temperature of annealing to attach properly to the template and let the polymerase incorporate the nucleotides. However, this implies that, when several of them are used in the same PCR reaction, a particular temperature that fits all the primers must be chosen for setting the Thermal Cycler conditions. To check for the appropriate temperature that provides with the maximum amplification, in both the Three PCRs Method and the QuikChange Site-directed Mutagenesis, several 25 μl of volume control reactions were performed with the primers needed in each round.

By employing a Thermal Cycler with Gradient option (MasterCycler, Eppendorf), PCR reactions were prepared with the primers to be tested and the rest of the reagents explained in section 3.1.3. Setting the instrument to give a different annealing temperature to each tube, the most suitable temperature was chosen by looking at the width and intensity of the amplification bands of each well in an agarose gel.

3.1.5. In-Fusion reaction

The resultant outcome from the Three PCRs Method (product PCR3) is a linear sequence with blunt ends at the 5' and 3' terminals. To clone this fragment into bacteria, it must be ligated to a linearized plasmid in order to get a circular plasmid that, among many genes, has inserted the mutated huDHOase sequence coming from the PCR3. The tool used to achieve that ligation was the In-Fusion HD cloning Kit (Clontech). In this kit, an enzyme called In-Fusion is in charge of fusing both the PCR3-generated fragment and the previously linearized vector, pOPIN-M. This enzyme recognized the flanking bases of the PCR3 fragment, as they are complementary to the ends of the linearized pOPIN-M (these bases were inserted by the primers DHOd and DHOr3 during the obtaining of PCR1 and PCR2), and joined both into a circular plasmid-like structure.

To prepare the In-Fusion reaction (PCR3 insert + pOPIN-M vector + In-Fusion Buffer 5x), Clontech's In-Fusion Calculator tool was used to determine the amount of vector and insert that was needed to obtain a high yield on the ligation. By inserting the size in base pairs for both of them, the amounts in nanograms that should be added are retrieved and for a 5 μl reaction, each reactant concentration was adjusted to match the recommendations from the manufacturer. After mixing the vector and insert in the In-Fusion Buffer 5x, the tube was kept at 50ºC for 15 minutes to let the enzyme perform its function.

As it will be explained in section 3.1.7, every time the In-Fusion reaction was done, a parallel control reaction with the same reagents except for the insert, which was substituted by milliQ water, was prepared. This control reaction was plated to make sure that the colonies that appeared in the In-Fusion petri dish did not emerge as a result of supercoiled plasmids [26] that did not insert the genetic material aimed during the In-Fusion reaction but, because they religated inside bacteria, they became resistant to the antibiotic of the plate and let colonies grow.

3.1.6. QuikChange Site-directed Mutagenesis

After the first round of mutagenesis in which the Three PCRs Method was used, the second and third rounds were done using a new strategy (Fig. 3.6) in which only one PCR is performed and an additional digestion step is needed. Generally, this method consist on:

1. Creating a so-called parental plasmid by ligating the template that will be used for mutation with a new plasmid, pCR^{TM}-Blunt.
2. Transformation of bacteria for the obtaining of methylated parental plasmids.

[26] twisted and tangled linear plasmids

3. Perform a PCR in which, after denaturation of the parental circular template, extension of the primers allow the synthesis of mutated nicked DNA.

4. Digestion of the remaining methylated parental non-mutated plasmid with DpnI so only mutated DNA is left.

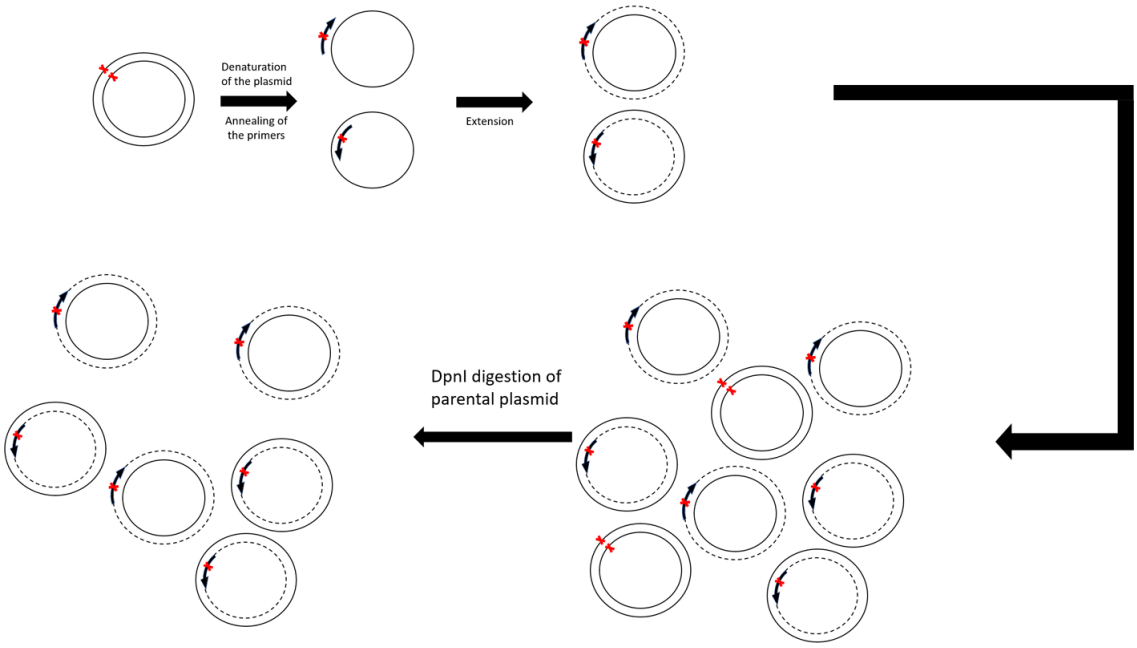


Fig. 3.6. QuikChange Site-directed Mutagenesis approach scheme.

Firstly, the PCR3 product coming from the first round, that now functions as template for the next mutagenesis, was ligated to the plasmid provided by the Kit, pCR^{TM}-Blunt (Fig. 3.7). To induce the ligation, a molar ratio of 1:3 vector pCR^{TM}-Blunt to insert (PCR3 fragment coming from round 1) was prepared, with the addition of the Ligation Buffer 10X and 0.5 units of T4 DNA ligase per μl of the final volume of the reaction.

To avoid the self ligation of the plasmid without the insert, the pCR^{TM}-Blunt was purchased already dephosphorilated. This dephosphorilation is needed as the T4 Ligase can only join the 3' hydroxyl group of one end with the 5' phosphate group of the other end. Therefore, for ligation to occur, at least one of the DNA ends (insert or vector) must contain a 5' phosphate group, in the case of this project the insert.

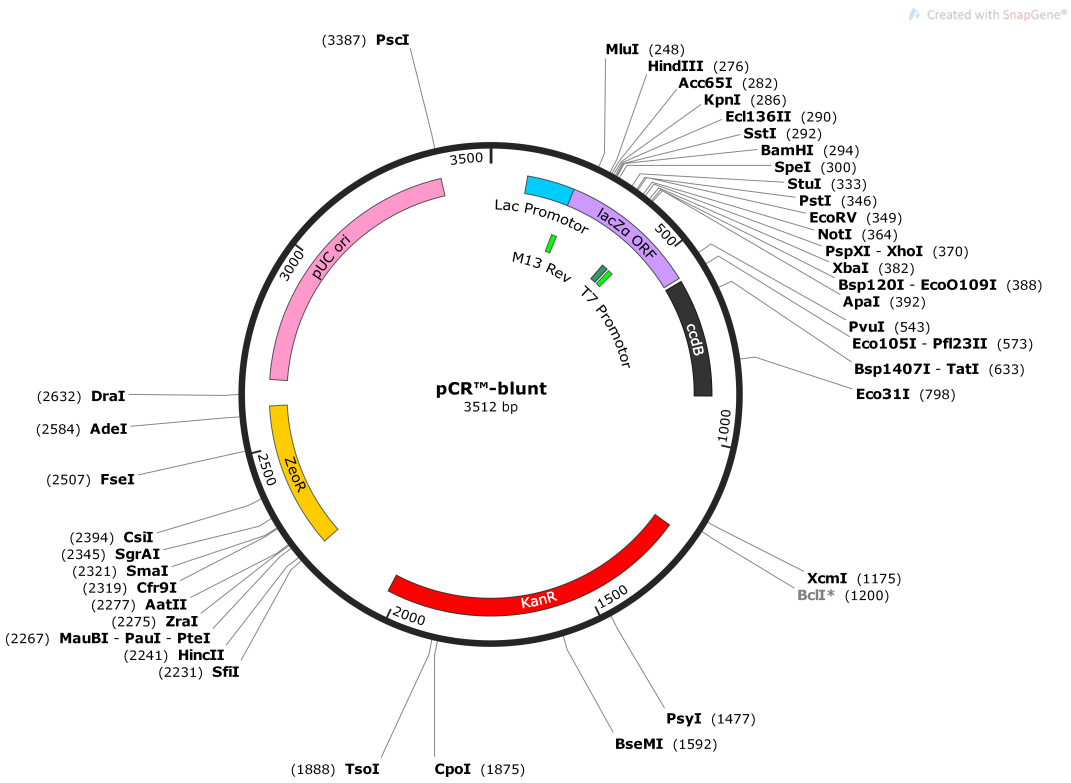


Fig. 3.7. Map of circular plasmid pCR^{TM}-Blunt.

After incubation for one hour at 16ºC, 4% of the ligation reaction was used for the transformation of TOP10 cells. In this step, methylation of the plasmid by bacteria is a key step as the parental non mutated plasmid must be eliminated at the end of the procedure. Moreover, the peculiarity about pCR^{TM}-Blunt is the presence of a lethal gene, ccdB, used to avoid false positives. This gene, inducing the death of the bacteria that has incorporated the plasmid, gets inactivated only when a DNA fragment is inserted in the middle of its sequence, preventing its expression. Therefore, colonies were able to grow in the plate due to the interruption of this gene by the PCR3 insert, a confirmation that all of them contained the proper ligation sought.

With the ligated construct ready and verified by Sanger sequencing, a single PCR with the corresponding pair of mutagenic primers and a proofreading DNA polymerase was done. In final volume of 20 μl the reaction consisted of: pCR^{TM}-Blunt ligated with PCR3, 2.5 μM of primer MR-mut2-d (Sigma), 2.5 μM of primer MR-mut2-r (Sigma), 10 mM dNTPs (ThermoFisher) , Buffer Exp.Long 10X (Sigma), 3.5 u Expand Long System Polymerase (Sigma) and 2% DMSO (ThermoFisher). The conditions for the thermal cycler were adapted to the new polymerase and for the annealing temperature, as the QuikChange method was used for the second and third rounds, the PCR was set at 68ºC in the annealing step as both pair of primers showed the greatest amplification with this temperature.

To eliminate only the methylated remnants of parental non-mutated plasmid in the reaction, digestion with 10 u of methylation-dependent endonuclease DpnI was done for an hour at 37ºC. After the treatment with DpnI, transformation of TOP10 was followed and the colonies obtained were inoculated to perform Minipreps[27] (GeneJET Plasmid Miniprep Kit, Thermo Fisher).

When the plasmid was isolated from bacteria, Sanger sequencing was used to verify that the sequence was properly mutated. The construct (pCR™-Blunt with huDHOase mutated up to round 2) was then used as template (no need for ligation step) to directly perform the site directed mutagenesis following the same procedure but this time using the primers corresponding to the third round (MR-mut3-d, MR-mut3-r).

3.1.7. Transformation of cells

In order to obtain a greater amount of methylated plasmids after using the In-Fusion technology or during the QuikChange Site-directed Mutagenesis, a transformation protocol was performed. In this protocol, the ability of *E.coli* bacteria to incorporate, methylate and replicate foreign external DNA material, was exploited to obtain exclusively the methylated plasmids that contained the mutated inserts desired for each round.

In this transformation procedure, aliquots of the TOP10 competent cell strain (Thermo Fisher) were gently mixed with 5-10% of the In-Fusion or QuikChange reactions and left on ice for 30 minutes in order for the bacterial outer membrane to stabilize. After that, the tubes were incubated in a water bath at 42ºC for 45 seconds in order for the heat to change the fluidity of the cell membrane and let the foreign genetic material to enter the cell. Immediately after the heat shock, the bacteria were put on ice for 3 minutes and then recovered with 300 μl of LB media culture for 1 hour at 37ºC in a shaker at 400 rpm.

To plate the bacteria, LB+Agar petri dishes were prepared with the addition of an antibiotic. In the case of the In-Fusion technology, every time transformation of cells was done after the procedure, Ampicillin 1000X was used. For the QuikChange reaction the antibiotic chosen was Kanamycin 1000X. The reason for the use of different antibiotics for each technology comes from the fact that the plasmid pOPIN-M contains in its sequence a gene that confers resistance to ampillicin, while pCR™-Blunt has a kanamycin one. In this way, only the bacteria that incorporates the plasmid into their cytoplasm acquires resistance to the antibiotic and is able to produce colonies in the antibiotic-containing agar plate. This is a common protocol to make sure that the colonies that appear in the petri dish have incorporated the genetic material aimed and no false posit-

[27] small scale procedure to isolate plasmid from bacteria

ives are used in further steps.

However, as supercoiled vectors can also knit their ends inside bacteria and express the antibiotic resistance gene, a control reaction against false positives, without insert and only vector, was always performed in the In-Fusion technology and plated along with the other In-Fusion reactions. In this plate, if no colonies were obtained, it implied that the vector used did not present supercoiling and the colonies that appeared in the other petri dishes plated with the In-Fusion reaction were all positive colonies with the insert+vector construction.

3.1.8. NaeI digestion test before sequencing

Before sending any plasmid for Sanger sequencing, Serial Cloner's Virtual Cutter tool (Serial Basics) was used first to find a restriction enzyme that could cut the plasmids, coming from each round of mutagenesis, in different sizes and fragments giving characteristic patterns. NaeI was the enzyme meeting the aforementioned requirement and digestion reactions were prepared at the end of each round. The concept behind the use of this enzyme is its ability as endonuclease to cleave DNA at distinct recognition sites. In this way, if the mutations performed at each round added or deleted any recognition site for NaeI in the sequence, then the enzyme would cut the DNA into more/less fragments with different sizes. As a consequence, by loading the digestion reaction (the plasmid cut by NaeI) into an agarose gel and comparing the number and size of the bands of the gel with the theoretical virtual cut, the mutations could be roughly confirmed before sequencing and no false positive would be sent.

Therefore, to exploit NaeI endonuclease function in the last three rounds of mutagenesis, digestion for one hour at 37ºC was done on reactions composed by 0.3 u/μl NaeI, CutSmart Buffer 10X (New England Biolabs) and 0.05 μg/μl plasmid (final product of each round of mutagenesis).

3.1.9. Preparation and analysis of sequencing results

To obtain confirmation on the step-by-step mutations, the plasmids of interest were prepared in a 200 ng/μl concentration and added to a pair of forward and reverse primers at 10 ng/μl. In this way, by elongation of both primers and the analysis by Sanger sequencing[28] on the resultant region of DNA, chromatograms were retrieved and evaluated to verify the mutations.

[28] technique used to determine the order of nucleotides of a DNA strand

For the visualization and verification of the Sanger sequencing results provided by the company Stab Vida, the chromatograms files were read with the 4peaks (Nucleobytes) software. Furthermore, to analytically compare amino acids and bases, and see if the sequence to be reached after each round was obtained, four virtual plasmid constructs with the huDHOase mutated according to each round were built with Serial Cloner (Serial Basics). These virtual mutated constructs were aligned to the Sanger sequencing results so the parity of bases and the presence of the mutations could be visually inspected and ease the confirmation of success.

3.2. Plasmid expression, protein production

To produce the protein on a large-scale, the final mutated construct addressed as pOPIN-M-huDHOase-ELF (ELF for: *E.coli* loop flexible) was expressed in a suspension of HEK293S-GnTI$^-$ cells (mammalian cells that perform suitable post-translational modifications) in Freestyle medium (Gibco) with 1% fetal bovine serum. For the transfection of these cells, the huDHOase-ELF plasmid was mixed with polyethylenimine (PEI: 25 kDa branched, Sigma), a polymer that binds to the DNA creating polyplexes and allows its delivery through the lipid bilayer, releasing it to the cytoplasm for its expression [28]. For the complexation, a 1:3 ratio mixture of DNA (1 μg/ml) and PEI was prepared by first diluting huDHOase-ELF to 20 μg/ml and PEI to 60 μg/ml in UltraDOMA medium (Lonza). Both dilutions were incubated for 5 minutes at room temperature and, after carefully mixing both solutions, another incubation step of 15 minutes was followed. The mix was then added to the cell culture and kept in an orbital stirrer at 125 rpm at 37ºC and 5% CO_2 for three days.

To test the transfection efficiency of the PEI used, plasmid pCMV-GFP (Addgene) containing the sequence for GPF[29] (Green Flourescent Protein) was complexated with PEI following the same protocol and transfected to a 25 ml sample of the cells. After three days, by using a hemocytometer and fluorescence microscopy on the sample (Leica Microscope DMLB Fluorescence and Phase Contrast), the amount of cells exhibiting green fluorescence were counted as well as the total number of cells on the four chambers of the hemocytometer. As the green fluorescence denotes the internalization of the polyplexes by the cells, the final efficiency of transfection was calculated as:

$$\text{PEI Transfection Efficiency} = \frac{\text{Number of cells that express GFP}}{\text{Total number of cells}} \times 100 \tag{3.1}$$

Afterwards, the cells transfected with pOPIN-M-huDHOase-ELF were harvested by centrifugation at 2500 rpm for 20 min and kept at -80ºC until the purification process.

[29] protein isolated from yellyfish that emits green fluorescence when exposed to UV light

3.3. Protein purification

To obtain the protein of interest huDHOase-ELF, two types of chromatography techniques were used (Fig. 3.9). Chromatography is used to separate components in a mixture by passing this mixture, contained in a mobile phase, through a stationary phase that interacts with the samples as they move past it. Depending on these types of interactions, the mobile phase can be partitioned in its elution and the samples of interest can be collected.

For the monitoring of the presence, purity, size and concentration of the protein, SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis) [29] and Bradford Assay [30] were used after each step of the purification.

3.3.1. Immobilized Metal Affinity Chromatography

The cell pellet was first resuspended in buffer NiA (20mM Tris pH 8, 500 mM NaCl, 5% Glycerol, 10 mM Imidazole, 2mM β-mercaptoethanol, 2mM phenylmethanesulfonyl-fluoride) and, by using a dounce homogenizer[30], the cells were subjected to shear in order to disrupt their membrane. To properly lysate the cells, the sample was sonicated by exposing it to pulses of 20 kHz for 15 seconds 5 times (all pulses followed by a resting period of 15 seconds) and kept on ice during all the process due to the possible overheating from the sonicator. Then, the lysate was centrifuged at 40,000 rpm for one hour at 4ºC and the supernatant obtained was filtered with a 0.45 μm membrane filter to avoid straining the Nickel Cartridge (ABT) [15].

As mentioned in section 3.1.1, the pOPIN-M-huDHOase construct used in this project was designed with several motifs in order to perform the transformation and purification steps. When the cell expressed the plasmid and produced the protein, as the plasmid had the 6xHis tag, MBP and PS site sequence immediately upwards to the huDHOase-ELF sequence, it attached these motifs to the N-terminal end of the huDHOase-ELF (Fig. 3.8) giving a recombinant protein with key elements to use for its purification.

Fig. 3.8. Graphical scheme of the sequence aimed to be expressed in HEK293S-GnTI$^-$cells. Before the sequence of huDHOase-ELF there is a tag consisting of six histidines, the sequence coding for the MBP and a site for cleavage by the PreScission protease.

[30]device used to aid the complete lysis of cells by tissue grinding

To first separate the huDHOase-ELF from the rest of the components of the lysate, immobilized metal affinity chromatography (IMAC) was used with a 5 ml Nickel Cartridge (Number 3 in Fig. 3.9) from ABT. Prior equilibration of the Nickel Cartridge with 10 volumes of buffer NiA, the lysate was loaded into the cartridge and, due to the high affinity of the 6xHis tag attached to the huDHOase-ELF to the nickel beads of the resin, the whole protein remained embedded in the cartridge adhered to the nickel.

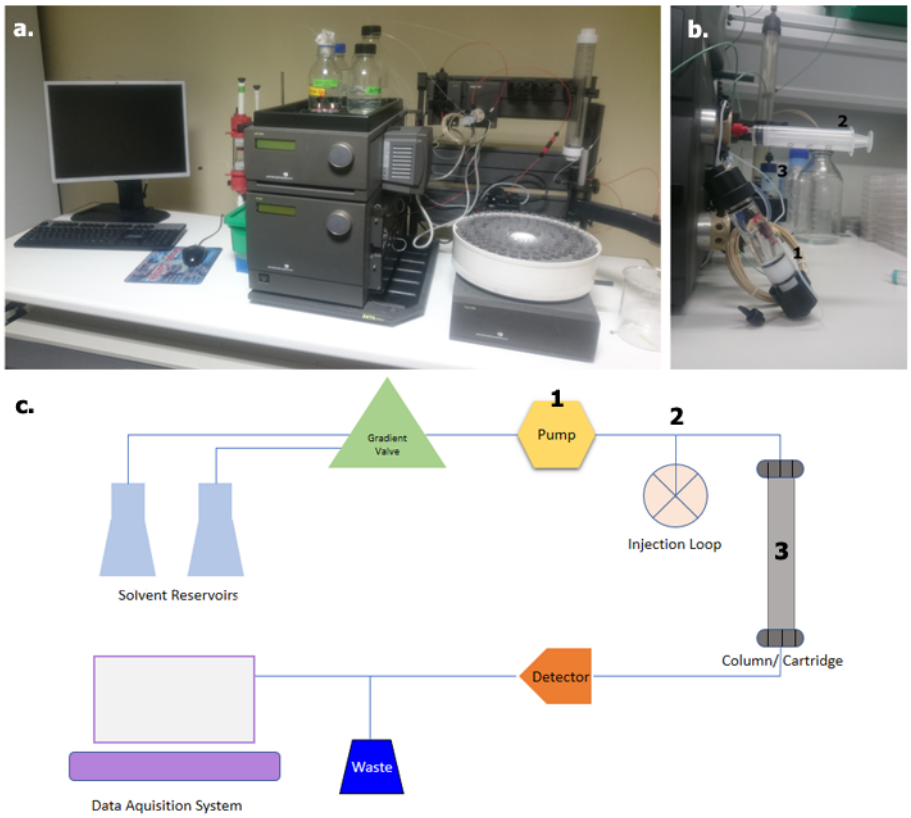


Fig. 3.9. FPLC (a.) and HPLC (b.) equipment used for the purification of huDHOase-ELF and their shared diagram (c.). Key parts of the HPLC system while doing IMAC are enumenrated. FPLC share the same components but with a vertical column instead of a cartridge.

The Nickel Cartridge was then connected to a high performance liquid chromatography system (ÄKTAprime plus HPLC, GE healthcare) and by washing it with 10% buffer NiB (20mM Tris pH 8, 500 mM NaCl, 5% Glycerol, 250 mM Imidazole, 2mM β-mercaptoethanol) the possible contaminants with low affinity to the nickel were removed. To elute the protein, 100% of buffer NiB (250 mM Imidazole) was injected into the cartridge so the high concentration of imidazole could compete against the protein for the nickel of the cartridge and detach it [15].

3.3.2. Dialysis and Digestion

To continue with the purification procedure, the concentration of imidazole in which the protein was eluted (250 mM) had to be reduced. To achieve this decrease of concentration, the protein was kept in a dialysis membrane (Sigma) against a dialysis buffer (20 mM Tris pH 8, 500 mM NaCl, 5% Glycerol, 30 mM Imidazole, 2mM β-mercaptoethanol) at 4ºC overnight.

To digest the protein and remove the 6xHis tag and the MBP, leaving only the huDHOase-ELF construct, PreScission protease that cleaves at the PS site (GE Healthcare) was added to the buffer NiA in a $1/20^{th}$ relation of the protein molecular weight and left overnight while dialysis took place. To extract the huDHOase-ELF protein, the nickel cartridge was connected again to the HPLC and, by injection of the digestion mixture and the dialysis buffer with 30 mM Imidazole, the huDHOase-ELF alone was recovered in the elution while the 6xHis tag and the MBP remained attached to the nickel beads of the cartridge's resin. To detach the 6xHis Tag + MBP from the resin, buffer NiB with 250 mM Imidazole was then injected into the cartridge.

3.3.3. Gel Filtration: Size Exclusion Chromatography

After concentration of the protein by centrifugation with the Amicon Ultra filters system (Sigma; 10 kDa cut off membrane), a gel filtration step was performed to purify the protein by size exclusion chromatography (SEC). For this protein, a Superdex 200 10/300 column was used (GE Healthcare) on a Fast Performance Liquid Chromatography system (ÄKTAFPLC, GE healthcare). The buffer prepared for the elution of the protein and the prior equilibration of the column consisted of 20mM Tris pH 8, 150 mM NaCl, 20 $ZnSO_4$ and 0.2 mM TCEP. The protein was loaded into the column and several fractions of volume (from 1 to 20) were retrieved. To determine in which fraction the protein of interest eluted, the Superdex 200 10/30 column calibration chromatogram provided by the manufacturer was used to extrapolate the peaks, corresponding to known molecular sizes, with the results obtained for the huDHOase-ELF. In this calibration test, several proteins with known molecular weight were passed through the column and eluted at different volumes depending on their size inducing different peaks of absorbance (measured at 280 nm). Therefore, with this test in consideration, the fractions of volume with the size corresponding to the elution of huDHOase-ELF were chosen.

To complete the purification procedure, the sample was concentrated using an Amicon Ultra filter system (Sigma; 5 kDa cut off membrane) and the final concentration of the protein was measured by Bradford Assay at 595 nm on a spectrophotometer (BioSpectrometer, Eppendorf) [30].

3.4. Activity assays on wt-huDHOase and huDHOase-ELF

To test the enzymatic activity of the huDHOase-ELF, activity assays of this protein and the wild type huDHOase were performed in parallel. Both proteins activities were assayed following the same procedure, the continuous measurement of absorbance from a spectrophotometer (BioSpectrometer, Eppendorf) at 230 nanometers, as at this wavelength the spectrophotometer can measure the concentration of the product dehydroorotate (DHO) present in the reaction [31].

For obtaining a reliable result, both directions of the enzymatic reaction: production of product DHO and production of substrate carbamoyl phosphate (CA-asp) (Fig 3.10), were reproduced in four essays that were performed in parallel on the chimeric huDHOase-ELF and wild-type huDHOase. In this way, the change on absorbance measured by the spectrophotometer in real time determined the ability of each enzyme to produce the product DHO from the substrate CA-asp (indicated by the increase of the absorbance measurements) or the use of DHO to produce the substrate CA-asp (by observing the decrease of the absorbance measurements). These four assays were conducted as follows:

I. Addition of CA-asp to wt-DHOase: retrieval of product production rate for the wild type huDHOase.

II. Addition of CA-asp to huDHOase-ELF: retrieval of product production rate for the chimeric huDHOase-ELF.

III. Addition of DHO to wt-huDHOase: retrieval of substrate production rate for the wild type huDHOase.

IV. Addition of DHO to huDHOase-ELF: retrieval of substrate production rate for the chimeric huDHOase-ELF.

	Protein to be tested	Direction of the reaction	Compound added to the reaction	Molecule produced in the reaction	Times the experiment was repeated
I	wt-DHOase	forward	CA-asp	DHO	3
II	huDHOase-ELF	forward	CA-asp	DHO	3
III	wt-DHOase	reverse	DHO	CA-asp	3
IV	huDHOase-ELF	reverse	DHO	CA-asp	3

TABLE 3.3. SUMMARY OF THE ACTIVITY EXPERIMENTS PERFORMED ON WT-HUDHOASE AND HUDHOASE-ELF

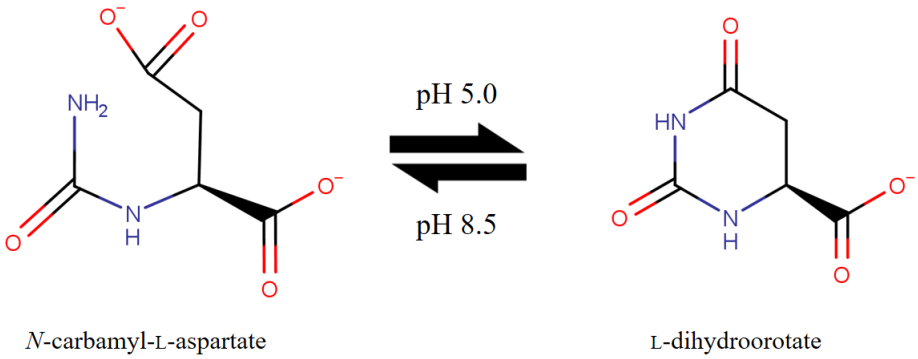


Fig. 3.10. Reversible cyclization reaction of CA-asp catalyzed by huDHOase and reproduced in the activity assays for wt-huDHOase and huDHOase-ELF.

To test the product production rate of both proteins, 10 μM wt-huDHOase was added to Activity Buffer A (50 mM NaPi pH 5.0, 150 mM NaCl, 0.1 mg/ml BSA and 20 μl $ZnSO_4$) in a final volume of 95 μl and 10 μM huDHOase-ELF was added in a separate tube to buffer A in a final volume of 95 μl. By filling two cuvettes with each reaction, the test started by taking 4-5 measures of the absorbance of the proteins alone without addition of the substrate. Once a stable measurement was obtained, 100 mM CA-asp was added to each cuvette and mixed fast by pipetting up and down for a few seconds while the spectrophotometer took an absorbance measure every 5 seconds for 5 minutes. The experiment was repeated three times for each protein and the data retrieved from the spectrophotometer's files were analyzed with GraphPad Prism.

For the evaluation on the substrate production rate favored by pH 8.5, 10 μM wt-huDHOase was added to Activity Buffer B (50 mM NaPi pH 8.5, 150 mM NaCl, 0.1 mg/ml BSA and 20 μl $ZnSO_4$) in a final volume of 95 μl while 10 μM huDHOase-ELF was added in a separate tube to buffer B in a final volume of 95 μl. Two cuvettes were filled with each reaction and, again, the test began with 4-5 measures of the protein absorbance alone. Once the measurements were stable, 10 mM DHO was added to each cuvette and rapidly mixed with a pipette for a short time while the spectrophotometer took measures of absorbance every 5 seconds for 5 minutes. As before, the experiment was repeated three times for each protein and the data retrieved from the spectrophotometer's files was analyzed with GraphPad Prism.

For the analysis of the data and the obtaining of the catalytic constant (k_{cat}) in each assay, the first few measurements (taken during the first seconds) were discarded as they were obtained while the mixing and handling of the device were performed. Therefore, to avoid this source of experimental error, all the outliers in the absorbance retrieved during those seconds were excluded.

To determine the activity of both proteins and compare them, the catalytic constant k_{cat} was computed. This constant k_{cat}, also called the turnover number, is a first order rate constant with reciprocal units of time that describes the limiting catalytic rate of an enzyme at saturation. Therefore, this k_{cat} represents the number of substrate molecules [S] transformed into product [P] in a given unit of time on a single enzyme [E] when this molecule is saturated with substrate [1].

$$E + S \underset{ES}{\rightleftharpoons} E + P$$

As the concentration of substrate added was three orders of magnitude higher than the concentration of enzyme, a steady state can be assumed. In this state, all the active sites of the enzyme are taking part in the [ES] complex and both rates of formation and breakdown of the [ES] complex are in equilibrium. Therefore, the k_{cat} can provide with the efficiency of the protein's catalytic behaviour in saturation.

$$k_{cat} = \frac{V_{max}}{[E]_t} \tag{3.2}$$

$$k_{cat} = \frac{\text{moles of DHO produced}}{\text{min} \cdot \text{concentration of protein added}} \tag{3.3}$$

A first plot was done to represent the absorbance measurements (in OD) at each time point in order to obtain the slopes (speed of the reaction) of each assay in OD/min. Then, to transform the measurements of absorbance into the concentration of DHO produced as a function of time (μM/min), the Beer-Lambert law was used [32]:

$$A = \varepsilon \cdot b \cdot C \tag{3.4}$$

Where A is the slope of the absorbance *vs.* time plot of each assay (OD/min), ε is the extinction coefficient[31] of DHO at 230 nm (1.17 mM^{-1} cm^{-1} [8]), b is the path length of the

[31] a measure of how strongly a substance absorbs light at a particular wavelength

light through the cuvette (1 cm) and C is the concentration of DHO per minute (M/min).

Finally, to obtain the k_{cat} of the wild type huDHOase and huDHOase-ELF in the two directions of the reaction, the concentration of DHO produced given by the Beer-Lambert law was divided by the concentration of protein added in the assay (0.5 μM) and, as each experiment was performed three times, the final four k_{cat} were given as an average between the three assays with their respective standard deviation.

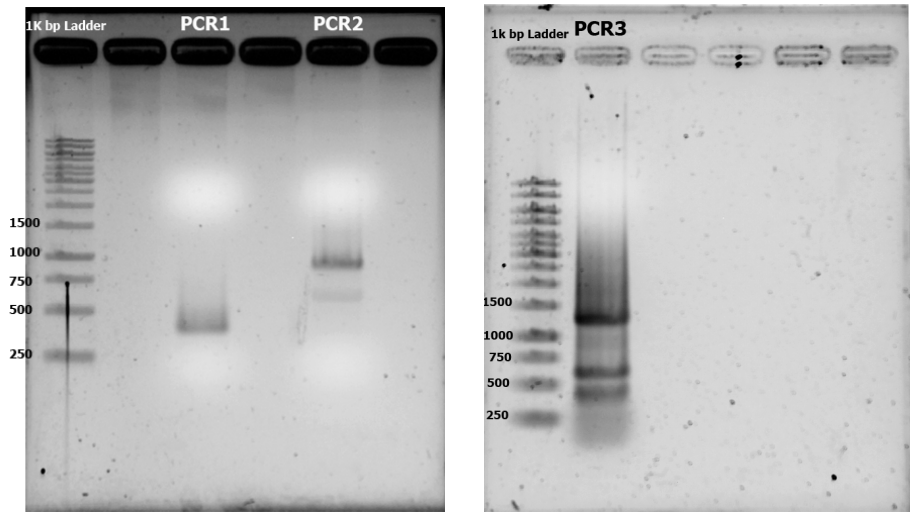
4. RESULTS

4.1. Cloning of huDHOase

4.1.1. Obtaining of PCR1 PCR2 and PCR3 fragments

For the first and fourth rounds of mutagenesis in which The Three PCR's Method was used, agarose gel electrophoresis gave a preliminary confirmation of the successful amplification of the DNA fragments by PCR. However, despite these bands being helpful to characterize the successful PCRs, they did not give any information about the presence of the aimed mutations. Further confirmation of the lengths and mutations of the fragments were obtained by chromatogram analysis after sequencing.

On the one hand, the 1.2% agarose gel in Fig. 4.1a shows how, for PCR1, a band appeared between the 250 and 500 bp size markers from the standard. This band corresponded to the 333 nucleotides fragment of huDHOase aimed to be mutated with the first three amino acids from *E.coli* flexible loop, plus 20 additional bases. These bases came from the DHOd flanking nucleotides that do not match the huDHOase sequence and were added to the sequence by elongation of the primer.



(a) Agarose gel loaded with the resultant PCR1 and PCR2 amplified by PCR for the 1st mutagenesis.

(b) Agarose gel loaded with the resultant PCR3 amplified by PCR for the 1st mutagenesis.

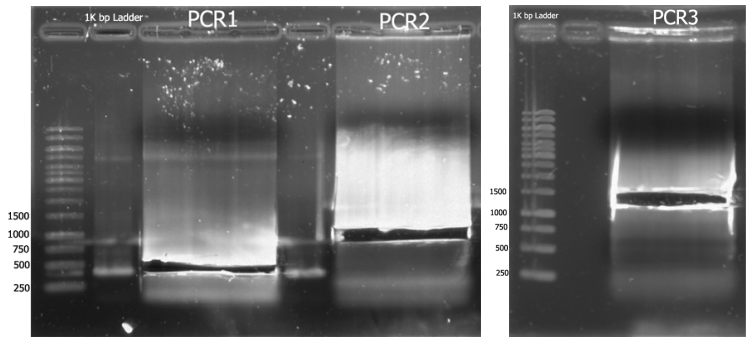
Fig. 4.1. Agarose gels providing with prior evidence on the obtaining of huDHOase fragments mutated according to the first round of mutagenesis.

Besides the PCR1 loaded well, another band corresponding to the loaded PCR2 reaction showed next to it. This band, between the 750 bp and 1000 bp size markers, coincided with the last 879 bases from the huDHOase also targeted to be mutated with the first three amino acids from *E.coli* flexible loop. Similarly to PCR1, to the previous 879 bases, 18 additional flanking nucleotides coming from the DHOr3 made a total of 897 nucleotides for the PCR2 construct.

On the other hand, as seen in Fig. 4.1b, the first time the PCR3 was performed, three distinct bands appeared in the gel. The first band, thicker and between the 1000 bp and 1500 bp size standards, coincided with the expected PCR3 fragment comprising 1211 nucleotides (the entire huDHOase sequence mutated with the first three amino acids from *E.coli* flexible loop plus the flanking bases from both DHOd and DHOr3 primers). Unexpectedly, below this band, two other blurry ones appeared. Further discussion on the presence of these two products can be found in Chapter 5.

Similar gels to Fig. 4.1 were obtained for the final fourth round of mutagenesis in which the fragments appeared between the same size markers although the sizes in this case were 383 bp for the PCR1, 876 bases for PCR2 and finally the whole mutated PCR3 fragment with the flanking bases from primers DHOd and DHOr3: 1217 bases. In the case of this fourth round of mutagenesis, no additional unexpected bands were obtained in the well corresponding to PCR3, but in PCR1 and PCR2. As before, refer to Chapter 5 for the discussion on this blurry bands appearing below the cuts.

In Fig. 4.2 a look on the resultant appearance of the excised gels from the 4th round is included. Special attention must be paid to the thin bands left on the wells of Fig. 4.2a.



(a) Agarose gel resultant from cutting both PCR1 and PCR2 bands of 4th mutagenesis.

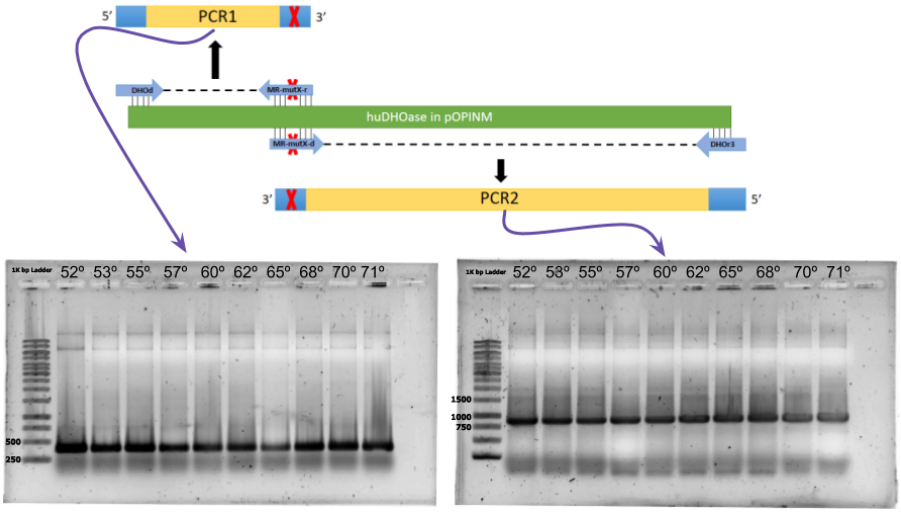
(b) Agarose gel resultant from cutting PCR3 band of 4th mutagenesis.

Fig. 4.2. Excised agarose gels showing the successful extraction of all three PCR products after the 4th round of mutagenesis for subsequent DNA purification.

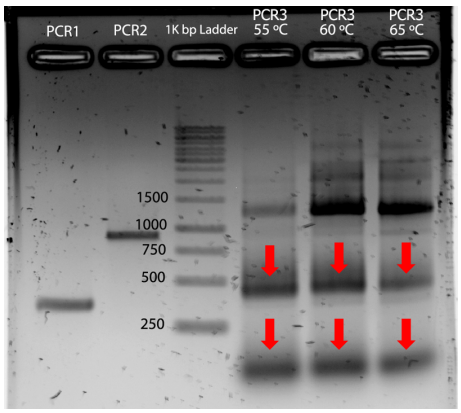
4.1.2. Temperature gradient tests

The gradient of annealing temperatures done to the primers from Table 3.2 revealed different sets of temperatures as the most amplifying–effective. On one side, all MR-mutX-r (the reverse primers for all rounds of mutagenesis) showed thicker and crisp bands exclusively from 52ºC to 55ºC reactions while for MR-mutX-d (the direct primers for all rounds of mutagenesis) irregular patterns were obtained in the gels, being 62ºC and 68ºC the only noteworthy temperatures showing a high amplification yield. An example on the appearance of the bands can be seen in Fig. 4.3a.

Regarding the DHOd and DHOr3 primers, in Fig. 4.3b, robust PCR3 bands were obtained only when the temperature exceeded 60ºC. Red arrows point towards unwanted PCR products, significantly of lower intensity at 65ºC of annealing.



(a) PCR1 and PCR2 done in a gradient of ten temperatures to test the amplification yield of MR-mut4-r (left gel) and MR-mut4-d (right gel).

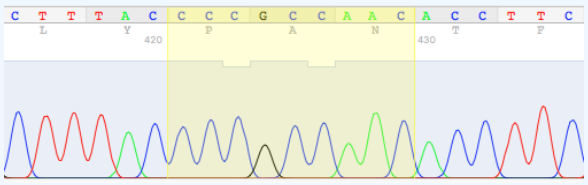


(b) PCR3 done in a gradient of three temperatures with DHOd/DHOr3.

Fig. 4.3. Gradient tests done for the Three PCRs Method in order to choose the appropriate annealing temperatures of the primers.

4.1.3. Ligation of huDHOase with pCR™-Blunt

After performing the ligation of pCR™-Blunt with the huDHOase sequence containing the mutations from the first cloning round, the chromatogram (Fig. 4.4a) retrieved from Stab Vida confirmed two facts: the mutation of the first three amino acids of the huDHOase flexible loop into the first three amino acids of the ecDHOase flexible loop, P A N: L1559P, N1560A, E1561N; the flanking bases before and after the huDHOase fragment corresponded to the pCR™-Blunt. With both mutation and ligation confirmed, the used of the QuikChange Site-directed Mutagenesis approach was approved to be continued.



(a) Chromatogram showing evidence of the first three amino acids mutated: P A N

```
                P  A  N  T  F  S  E  L  R  L  D  S  V  V  Q  W  M  E  H  F
Seq_1     ACCCCGCCAACACCTTCTCTGAGCTGCGGCTGGACAGCGTGGTCCAGTGGATGGAGCATT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2     accccgccaacaccttctctgagctgcggctggacagcgtggtccagtggatggagcatt
                P  A  N  T  F  S  E  L  R  L  D  S  V  V  Q  W  M  E  H  F
```

(b) Alignment of the theoretical construction made in Serial Cloner (Seq_1) with the sequencing file (Seq_2) from Stab Vida. Bases that comprise the first round of mutagenesis are highlighted in green.

Fig. 4.4. Results obtained by Sanger sequencing in which both the ligation of huDHOase with pCR™-Blunt and the first round of mutagenesis are verified.

4.1.4. Digestion with NaeI

The use of NaeI to digest the plasmids before sending them for Sanger sequencing, resulted in gels like the ones showed in Fig. 4.5 and Fig.4.6. As seen in Fig.4.5, the loaded pCR™-Blunt plasmid mutated up to the second round evidenced the accuracy of Serial Cloner's Virtual Cutter as the four characteristic bands given by the software were experimentally obtained. The next 8 wells, loaded with 8 plasmids samples coming from different colonies after the third round of mutagenesis, all showed an identical pattern that matched the theoretical cut from the construction made with Serial Cloner. Therefore, as the digestion showed in an indirect way that the huDHOase of the plasmids was mutated as intended, all eight plasmids were considered suitable for verification though sequencing.

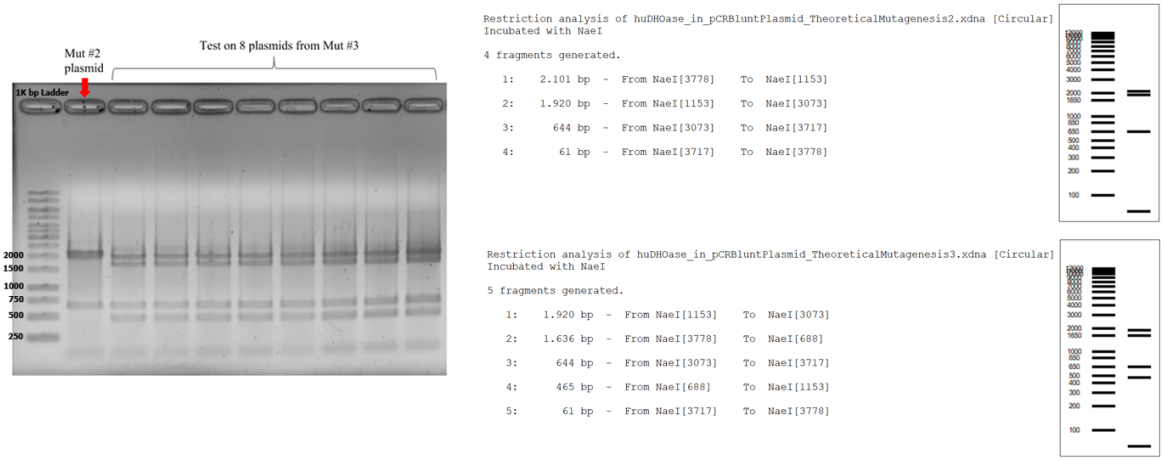


Fig. 4.5. Agarose gel loaded with NaeI-digested plasmids from the 2nd and 3rd rounds of mutagenesis next to a graphic report from Serial Cloner's Virtual Cutter. As the bands of the gel followed the pattern dictated by the Virtual Cutter, the mutations planned on both plasmids were considered accomplished.

Similarly, in Fig. 4.6, the bands of the gel showed how the enzyme digested differently the plasmid containing the huDHOase sequence mutated up to the third round (previously verified through sequencing) and the plasmid isolated from a colony from the fourth round. And, at the same time, they matched with the theoretical patterns expected from the virtual cut done. Thus, the three bands were considered proper evidence of the presence of the mutations and the cloning procedure was foreseen as terminated pending for the final corroboration from Sanger sequencing.

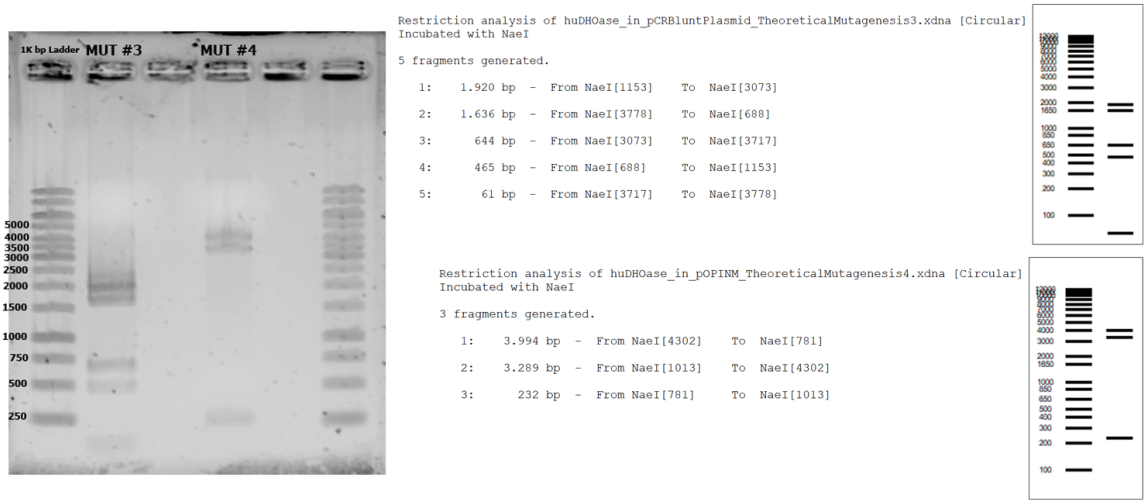


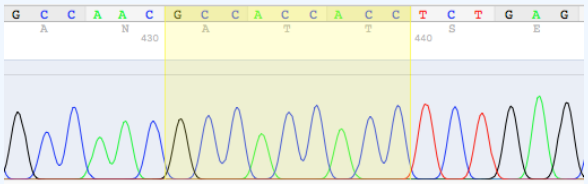
Fig. 4.6. Agarose gel loaded with NaeI-digested plasmids from the 3rd and 4th rounds of mutagenesis next to a graphic report from Serial Cloner's Virtual Cutter. As both the experimental and theoretically expected bands from Serial Cloner were obtained in an exact arrangement, the mutation was considered successful.

4.1.5. Sequencing results of the huDHOase-ELF

After each round was completed by the Three PCRs Method or the QuikChange Site-directed Mutagenesis approach, the analysis of the chromatograms and the resultant sequencing files determined the progressive success on mutating huDHOase flexible loop amino acids with the corresponding amino acids from ecDHOase flexible loop:

Human DHOase flexible loop	L N E – T F S E L – R L D
E.coli DHOase flexible loop	P A N A T T N S S H G V T

As shown in Fig. 4.7, the chromatogram ABI file retrieved shows: the insertion of one alanine, N1561_T1562insA, and the mutation of a phenylalanine into a threonine, F1564T. The bases comprising both the insertion and the substitution corresponded to the bases of the designed primers and thus, the ones aimed.



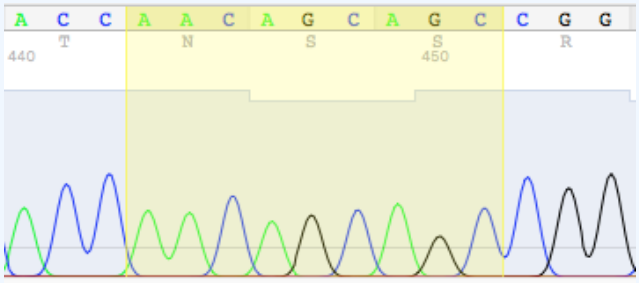
(a) Chromatogram confirming the insertion of an Alanine and the substitution of huDHOase F into ecDHOase T

```
                P  A  N  A  T  T  S  E  L  R  L  D  S  V  V  Q  W  M  E  H
Seq_1       ACCCCGCCAACgccaccaccTCTGAGCTGCGGCTGGACAGCGTGGTCCAGTGGATGGAGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2       accccgccaacgccaccacctctgagctgcggctggacagcgtggtccagtggatggagc
                P  A  N  A  T  T  S  E  L  R  L  D  S  V  V  Q  W  M  E  H
```

(b) Alignment of the theoretical construction made in Serial Cloner (Seq_1) with the sequencing file (Seq_2) from Stab Vida. Bases that comprise the 2nd round of mutagenesis are highlighted in light green while past mutations are marked in dark green.

Fig. 4.7. Results obtained by Sanger sequencing of the pCR™-Blunt vector ligated with huDHOase in which confirmation of the insertion and mutation aimed for in the 2nd round of mutagenesis was obtained.

In Fig. 4.8 is presented the reading of the third chromatogram on 4peaks and the alignment on Serial Cloner with the theoretically expected construct for the third round. Both show three substitutions in the huDHOase: S1565N, E1566S and L1567S (Fig. 4.8a) immediately after the rest of changed bases from the past 1st and 2nd rounds that remained unaltered (Fig. 4.8b).



(a) Chromatogram verifying the mutation of huDHOase S E L to ecDHOase N S S

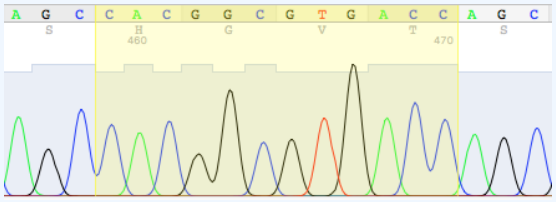
```
                P  A  N  A  T  T  N  S  S  R  L  D  S  V  V  Q  W  M  E  H
Seq_1           ACCCCGCCAACGCCACCACCAACAGCAGCCGGCTGGACAGCGTGGTCCAGTGGATGGAGC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2           accccgccaacgccaccaccaacagcagccggctggacagcgtggtccagtggatggagc
                P  A  N  A  T  T  N  S  S  R  L  D  S  V  V  Q  W  M  E  H
```

(b) Alignment of the theoretical construction made in Serial Cloner for the 3rd round (Seq_1) with the sequencing file (Seq_2) from Stab Vida. Target bases that comprised the 3rd round of mutagenesis are highlighted in light green while past mutations and insertions are marked in dark green.

Fig. 4.8. Results obtained by Sanger sequencing of the pCR™-Blunt vector ligated with huDHOase in which confirmation of the mutations aimed in the 3rd round of mutagenesis was obtained.

In the case of the fourth round, as can be seen in Fig. 4.9b: (1) three bases coding for histidine were inserted (S1567_G1568insH) between the past mutated Serine and a Glycine; (2) the last amino acids of the huDHOase flexible loop R L D were substituted into G V T (R1569G, L1570V, D1571T); (3) an unexpected and unwanted base substitution appeared, g.4695C>T, but the resulting codon[32] translates into the same amino acid, an asparagine on position 1565.

[32] sequence of three adjacent DNA bases that code for an amino acid when translation into protein occurs, one of the *rules* that define the genetic code



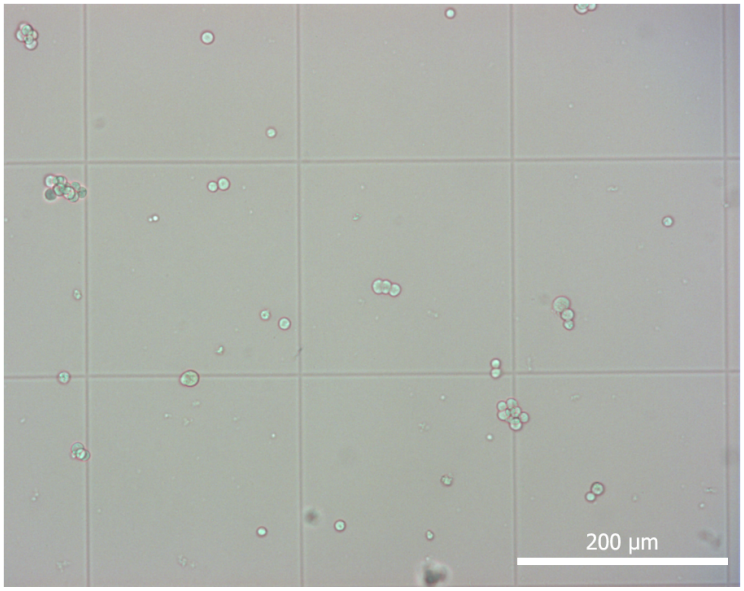
(a) Chromatogram showing the insertion of three bases coding for histidine and the mutations of the huDHOase sequence: R L D into the ecDHOase sequence: G V T

(b) Alignment of the theoretical construction made in Serial Cloner for the 4th round (Seq_1) with the sequencing file provided by Stab Vida (Seq_2). Bases inserted and mutated in the 4th round are highlighted in light green while past mutations and insertions are marked in dark green. The unwanted base substitution is circled in orange.

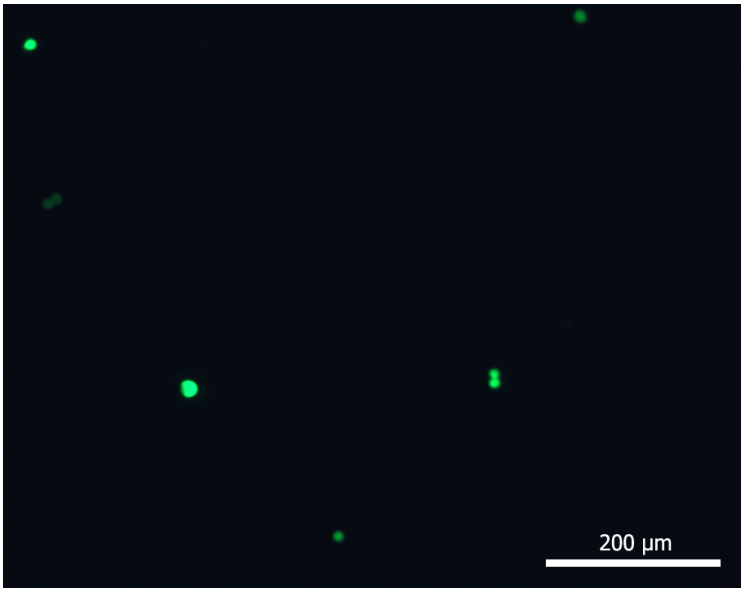
Fig. 4.9. Results obtained by Sanger sequencing of the pOPIN-M vector ligated with huDHOase in which confirmation of the insertion and mutations aimed in the 4th round of mutagenesis was obtained.

4.2. Plasmid expression, protein production

After three days, 2.6 million of cells were obtained for the culture in which polyplexes of huDHOase-ELF + PEI were transfected. Regarding the 25 ml control culture transfected with the GFP contaning plasmid (pCMV-GFP), 21% efficiency was observed by fluorescence microscopy. An image of the cells while being counted on a hemocytometer is shown on Fig. 4.10a while the production of the GFP protein and its expression of green fluorescence can be seen on Fig. 4.10b.



(a) HEK293S-GnTI$^-$ under the light microscope after being transfected with plasmid pCMV-GFP.



(b) HEK293S-GnTI$^-$ under fluorescent microscopy after being transfected with GFP containing plasmid pCMV-GFP.

Fig. 4.10. Analysis on HEK293S-GnTI$^-$ transfection efficiency by fluorescence microscopy.

4.3. Purification of huDHOase-ELF

4.3.1. Elution of the histidine tagged huDHOase-ELF bound to MBP after IMAC

As seen on Fig. 4.11, the chromatogram obtained from the HPLC system shows, as different peaks of conductivity, both the discard of contaminants and the elution of the histidine tagged huDHOase-ELF, still attached to MBP, in different fractions when IMAC was performed.

After the wash with 10% NiB, a first conductivity peak of approx. 280 mAU appeared 10 minutes into the experiment. The eluted elements, taken since the growth of the peak to its decline, were pooled in fractions 1, 2 and 3 and considered as contaminants. The third 100 mAU peak, retrieved 40 minutes into the procedure and obtained after the injection of 100% NiB to the cartridge, corresponds to the elution of the huDHOase-ELF (along with other contaminants) which were collected from fraction 4 to 6.

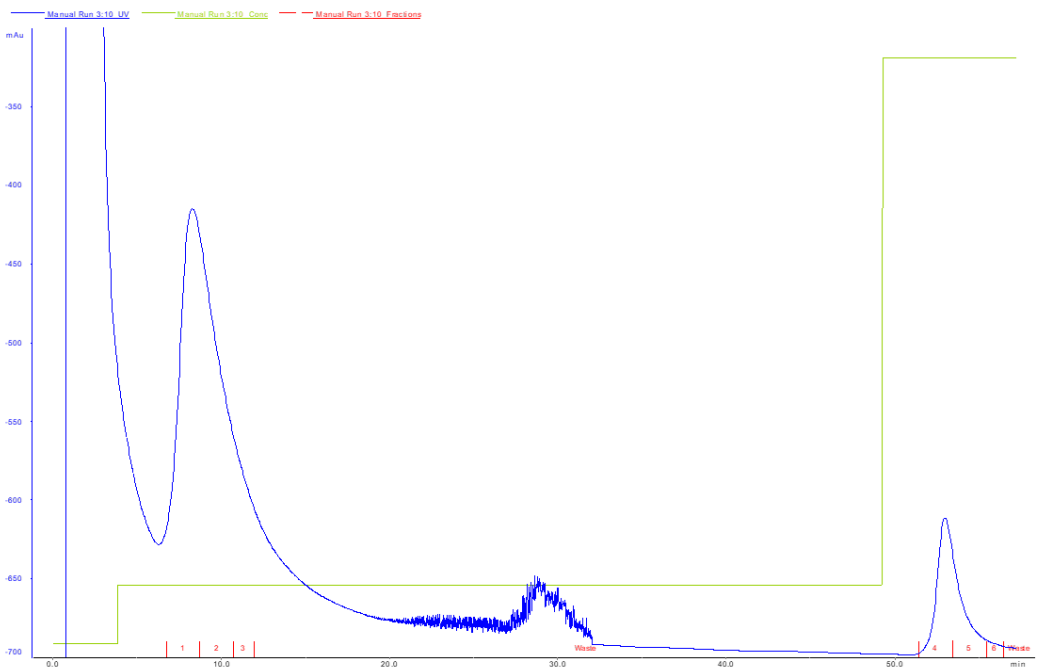


Fig. 4.11. Graph retrieved from the HPLC system during IMAC. Contaminants correspond to fractions 1 to 3 where a 280 mAU peak is observed. The 100 mAU of conductivity peak corresponds to the elution of the protein of interest along with other contaminants in fractions 4, 5 and 6.

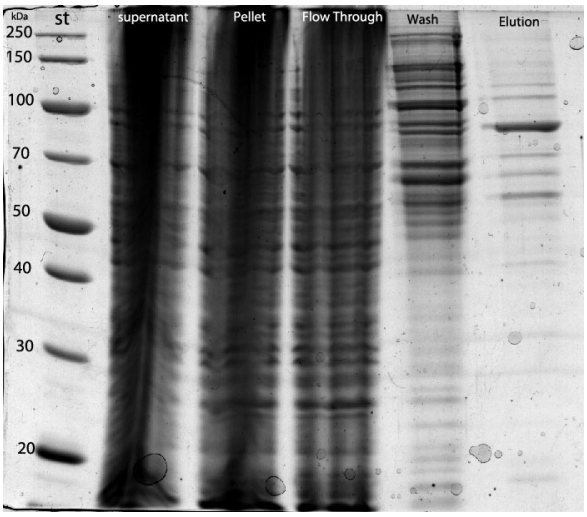


Fig. 4.12. SDS-PAGE gel loaded with a molecular size marker (st), the supernatant taken after centrifugation of the cells, a sample of the pellet, the flow through after loading the cartridge with the lysate, a contaminant sample from fraction 2 and the elution of the protein from fraction 5.

Fig. 4.12 shows the SDS-PAGE gel prepared after the obtaining of the previous results. The elution well, corresponding to a sample coming from fraction 5, showed the 6xHis Tag + MBP + huDHOase-ELF between the 70 and 100 kDa size marks, along with many contaminants of different size and abundance.

Up to this point, by measuring with Bradford assay [30], the concentration of the protein construct that eluted in 13 ml from fractions 4-6 gave 0.3 mg/ml.

4.3.2. Detachment of the 6xHis Tag and MBP through digestion

When the digested and dialyzed protein sample was loaded again into the cartridge, the connected HPLC system returned the graphical tracking shown on Fig. 4.13. By injection of the dialysis buffer, the elutions corresponding to the first conductivity increment, close to 40 mAU, were returned in 12 fractions among which the huDHOase-ELF was retrieved. When 100% NiB was passed through the column, a second conductivity peak, almost at 50 mAU, represented the pool of the 6xHis Tag and MBP on fractions from 13 to 19.

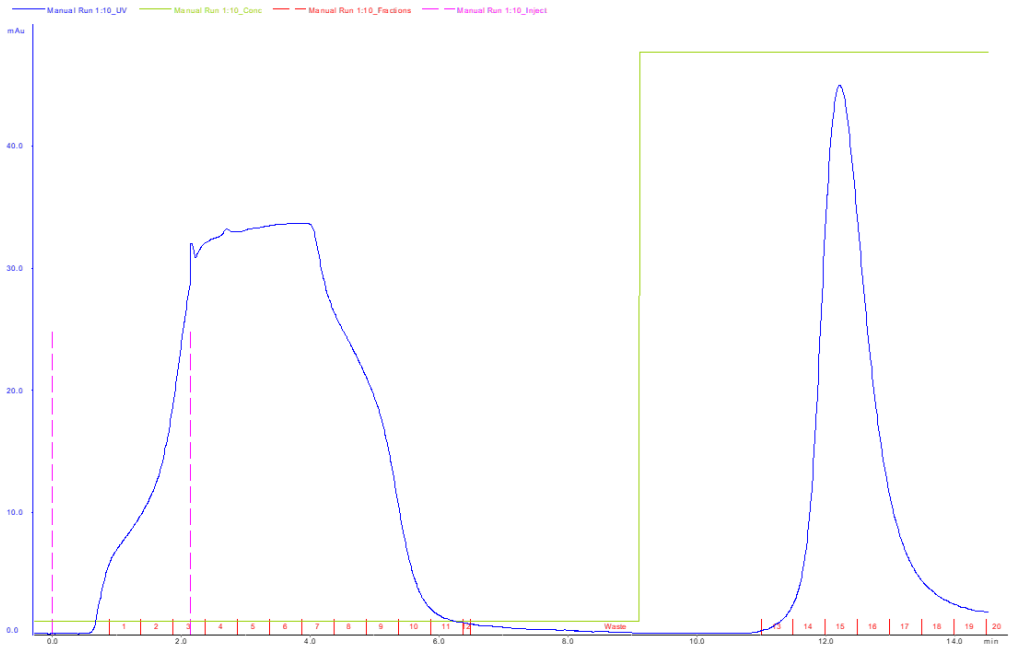


Fig. 4.13. Chromatogram retrieved from the HPLC system showing the elution of the untagged protein from fractions 1 to 12 in the 40 mAU peak, and the pool of the 6xHis Tag and MBP on fractions from 13 to 19 almost at 50 mAU.

By looking at the SDS-PAGE gel on Fig. 4.14, the sample of huDHOase-ELF + 6xHis Tag + MBP before digestion served as control, as again migrated between the 70 and 100 kDa markers of size. On its third well, a digested sample before its injection to the cartridge for separation was loaded. In this case, three bands very close to each other showed on the gel indicating the decoupling aimed. Fractions 2, 7, 9 and 11 ran above the 40 kDa marker showing, unexpectedly, two bands. The bands which migrated from wells 13, 14, 15 and 16 confirmed the elution of the 6xHis Tag + MBP (approx. 40 kDa) and their disassociation from the huDHOase-ELF.

After checking the SDS-PAGE gel, concentration by centrifugation was followed and the 6 ml pool of fractions 14-17 was finally retrieved in a concentration of 0.2 mg/ml.

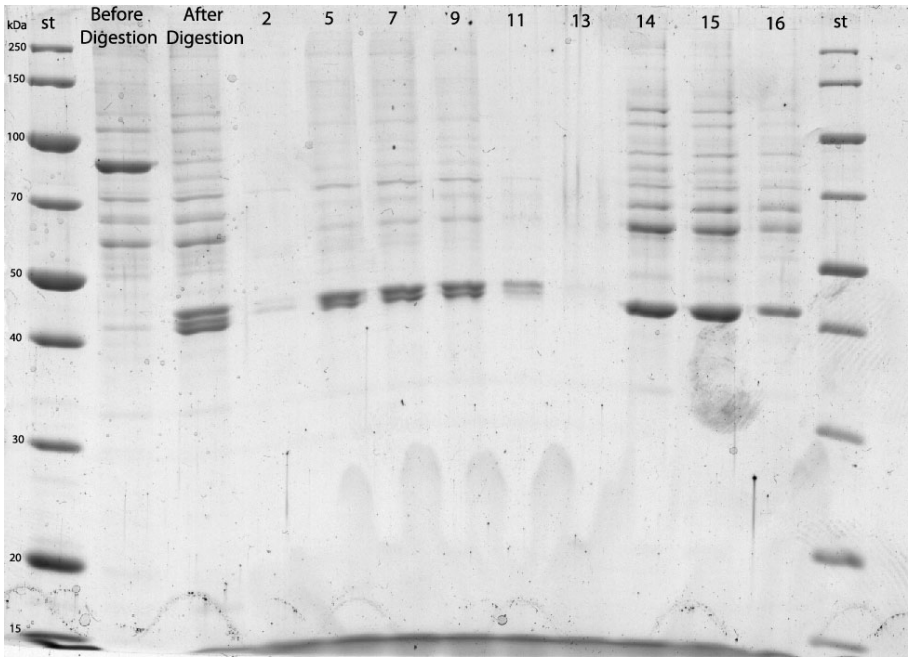


Fig. 4.14. SDS-PAGE gel with both molecular size markers in the first and last well. The controls here used were the samples of the protein before and after digestion taken before loading into the HPLC system. The wells corresponding to fractions from 2 to 11 present all the expected band between 40-50 kDa corresponding to the huDHOase-ELF along with a very close secondary band underneath it. The 6xHisTag + MBP can be seen on wells 14-16 where a band, thicker and lower than the huDHOase, ran between the 40 and 50 kDa size markers.

4.3.3. Purified huDHOase-ELF after SEC

When size exclusion chromatography was performed in the last step of the purification, after loading the untagged protein to the Superdex 200 10/300 column, a peak of absorbance of almost 30 mAU first appeared when the 14-16 fractions were being collected (Fig. 4.15). According to the manufacturer's chromatograph test, the eluted samples between these two fractions are between 35 and 67 kDa in size. Therefore, the pool starting at the growth of the peak and finishing at its decline, that was eluted from fractions 12 to 21, was set apart. The rest of the contaminants, smaller in size, were eluted from fractions 21 to 33 following the irregular peak depicted in the graph.

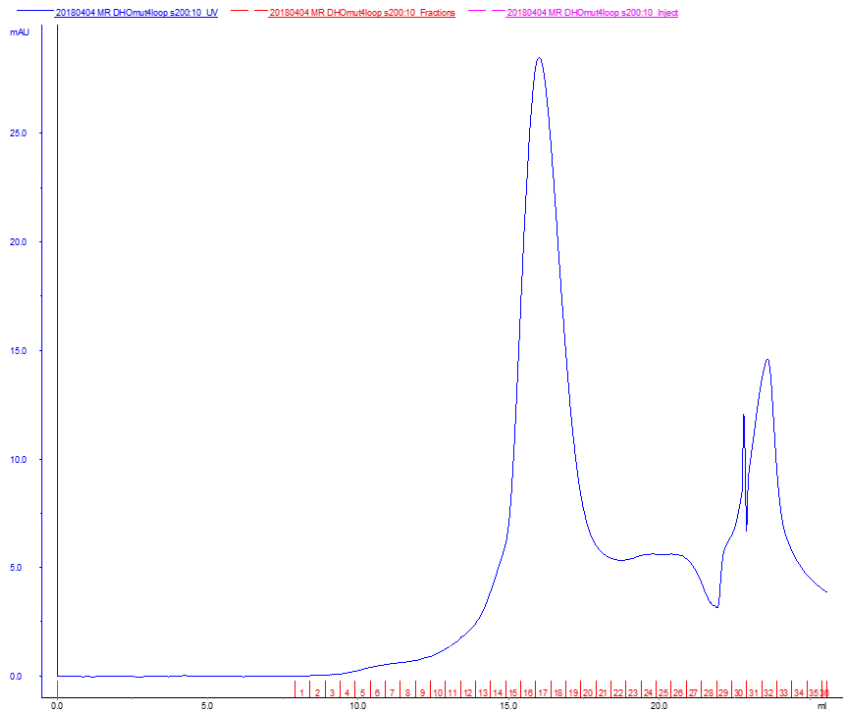


Fig. 4.15. Chromatogram retrieved after SEC. The expected huDHOase-ELF eluted during the peak at 30 mAU, between fraction 14 and 16. The samples that were obtained from fractions 21 to 33, and induced irregular peaks of conductivity, were considered contaminants.

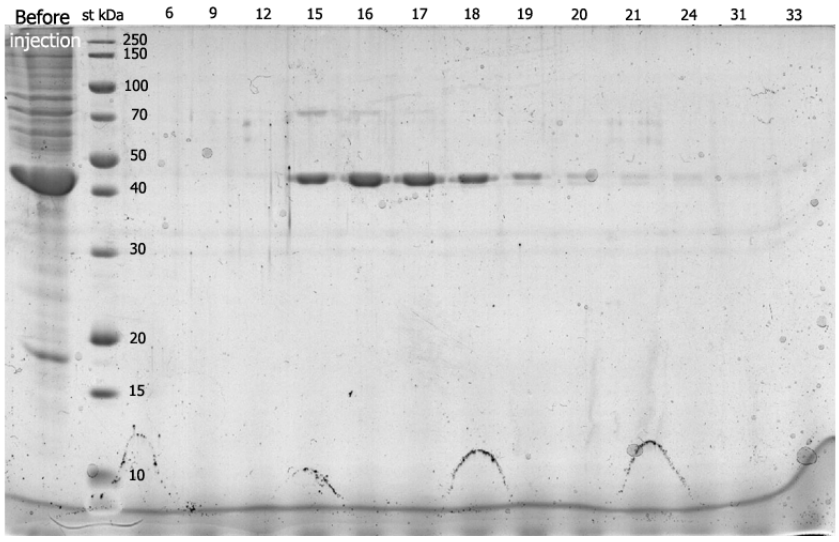


Fig. 4.16. SDS-PAGE gel obtained after performing SEC on the protein. A sample before injection shows the protein plus other contaminants in the first well. A molecular size marker separates this well from the fractions obtained by size exclusion chromatography (from 6 to 33). From 15 to 24 the wells present a visible band between 40 and 50 kDa equivalent to the purified huDHOase-ELF.

On Fig. 4.16, the SDS-PAGE gel loaded with the fractions coming from the Superdex 200 10/300 column elutions is shown. As control, a sample of the protein before injection was run showing the protein plus contaminants and other unwanted residues. From the 15 to 19 fractions wells, 5 easily distinguishable bands can be observed between the 40 and 50 kDa size markers. And, as the bands started fading with the progress of the fractions, this was considered the sought huDHOase-ELF.

Lastly, the pool from fractions 14 to 21 was concentrated by centrifugation and 1.2 mg/ml huDHOase-ELF was obtained in 60 μl.

4.4. Catalytic activity of huDHOase-ELF

For the forward reaction: CA-asp $\xrightarrow{\text{pH 5}}$ DHO, wt-huDHOase presented a k_{cat} of 113.5 ± 0.5 min^{-1} while for huDHOase-ELF k_{cat}=1.7 ± 1.1 min^{-1}. Thus, in this direction, the catalytic rate of the wt-huDHOase is 67 times higher than the one presented by the chimera huDHOase-ELF.

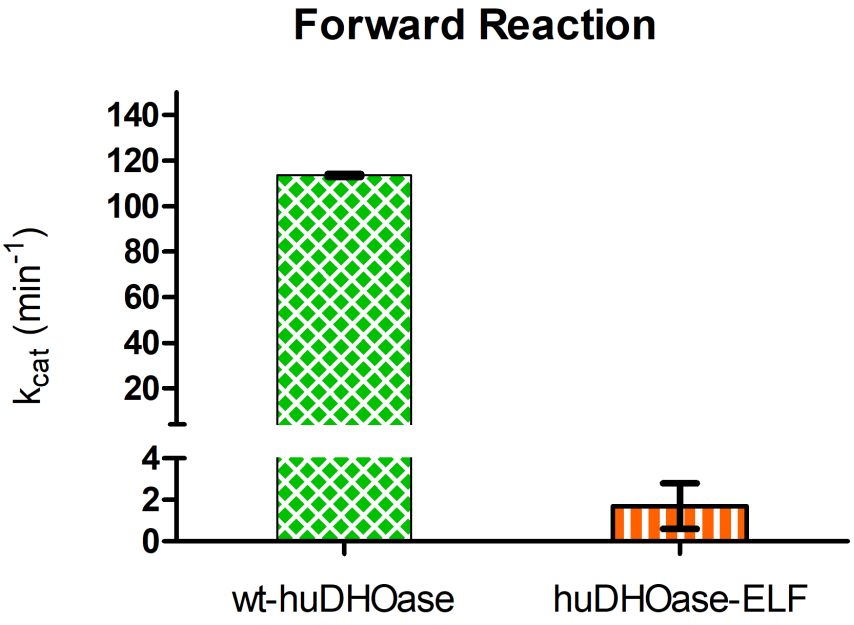
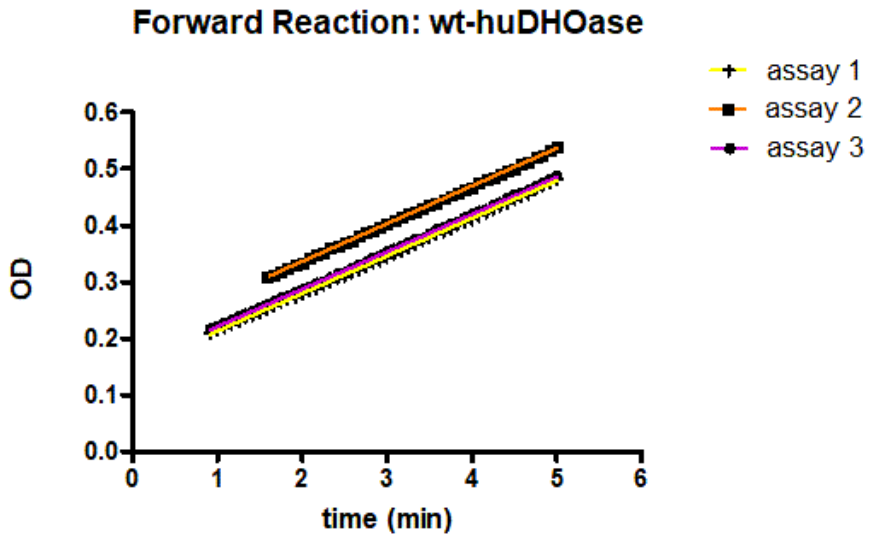
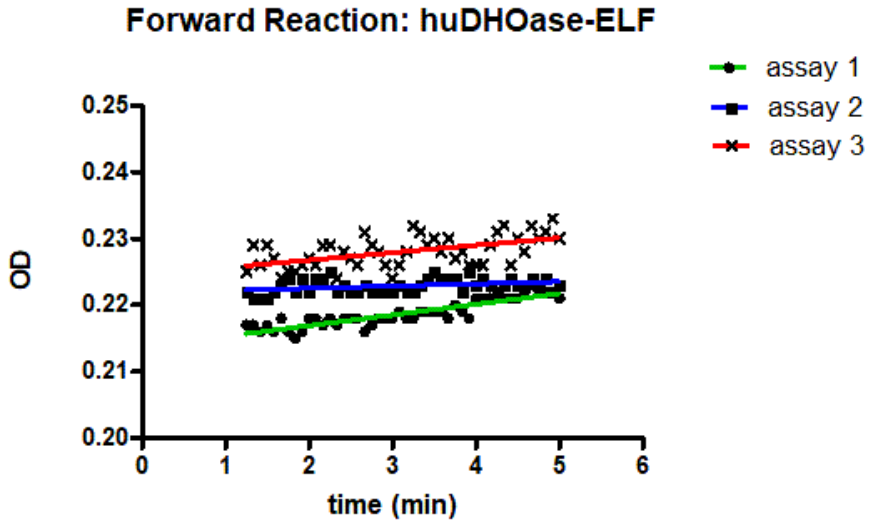


Fig. 4.17. k_{cat} values for wt-huDHOase and huDHOase-ELF in the forward direction of the reaction.

As seen in Fig. 4.18, the absorbance measurements taken by the spectrophotometer present a positive slope following the progressive production of the product DHO as it absorbs at 230 nm. While wt-huDHOase presented a distinct and well-defined increase of absorbance, huDHOase-ELF assays were irregular, showing oscillations in the measurements (e.g. huDHOase-ELF assay 3 on Fig. 4.18b).



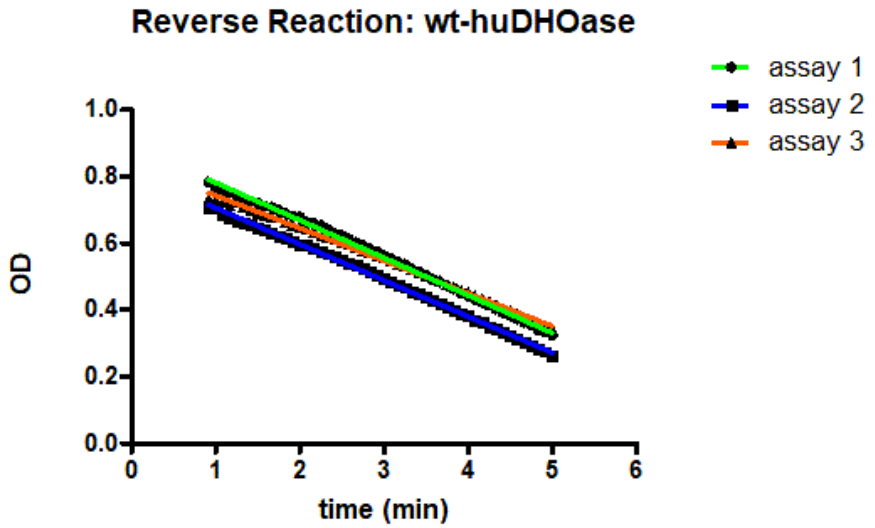
(a) Absorbance measurements for 5 min at 230 nm when CA-asp was added to the wt-huDHOase for the forward reaction to take place.



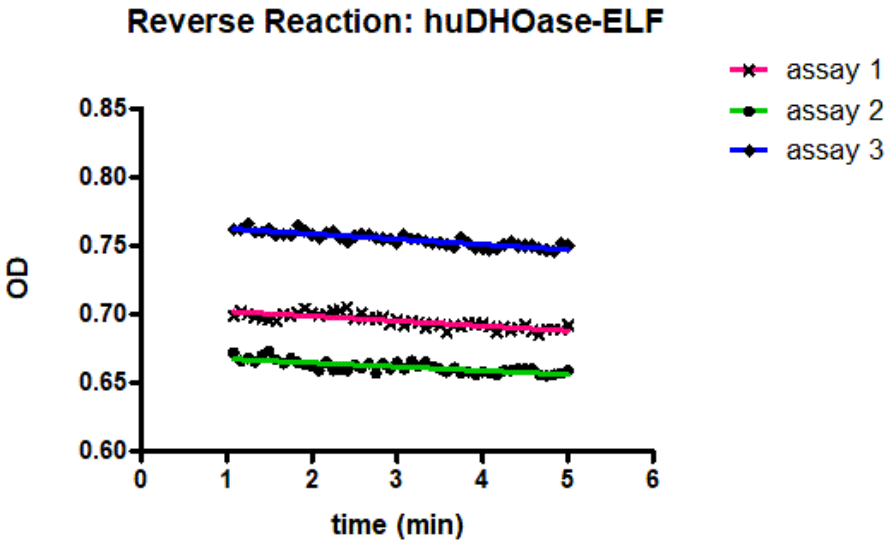
(b) Absorbance measurements for 5 min at 230 nm when CA-asp was added to the huDHOase-ELF for the forward reaction to take place.

Fig. 4.18. Graphs representing the increase in the absorbance (measured at 230 nm) when DHO was being produced by both wt-huDHOase and huDHOase-ELF for five minutes.

When the last three assays were performed on both wt-huDHOase and huDHOase-ELF, the spectrophotometer measurements indicated the gradual depletion of DHO as it was being transformed into the substrate CA-asp, giving negative slopes for both reverse reactions.



(a) Absorbance measurements for 5 min at 230 nm when DHO was added to the wt-huDHOase for the reverse reaction to take place.



(b) Absorbance measurements at 230 nm for 5 min when CA-asp was added to the huDHOase-ELF for the reverse reaction to take place.

Fig. 4.19. Graphs depicting the reverse reaction measured at 230 nm: the consumption of DHO as CA-asp was being produced by both wt-huDHOase and huDHOase-ELF.

For the reverse reaction: CA-asp $\xleftarrow[\text{pH 8.5}]{}$ DHO, wt-huDHOase presented a k_{cat} of 181.8 ± 13.2 min^{-1} while for huDHOase-ELF k_{cat}=5.9 ± 0.9 min^{-1}. In this direction, wild type huDHOase exhibited a 31-fold higher rate in comparison with the mutant, huDHOase-ELF.

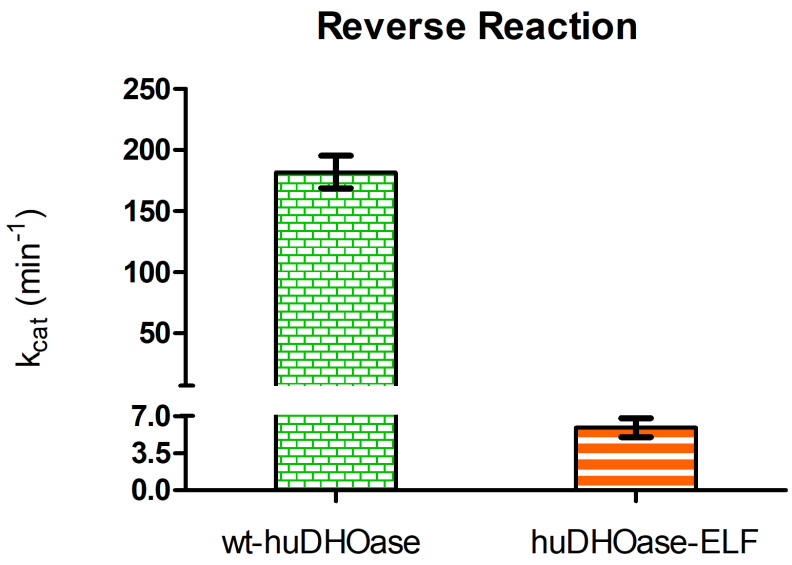


Fig. 4.20. k_{cat} values for wt-huDHOase and huDHOase-ELF in the reverse direction of the reaction.

5. DISCUSSION

Advantages and disadvantages behind the cloning approach adopted

The main reason behind the design of a different and atypical protocol for the cloning of the huDHOase is the plasmid pOPIN-M in which is inserted. As the huDHOase cannot be expressed alone, but with several elements crucial for its purification: 6xHis tag, MBP and PreScission site; a novel construction of pOPIN-M with the additional elements was needed so the cells transfected with the plasmid could also express them. However, by having the huDHOase sequence in this plasmid, the use of the common mutagenesis protocols that force the use of other vectors (like the pCR^{TM}-Blunt in the case of the QuikChange Site-directed Mutagenesis from Agilent used in other published works [11]) is hindered. Therefore, the initial and final application of the Three PCRs Method for the first and fourth round of mutagenesis, demonstrated to be a robust strategy to extract the huDHOase sequence from the pOPIN-M-huDHOase vector (1st round) and then insert it again so the purification elements could be expressed attached to the protein (4th round). And, at the same time this was done, the Three PCRs Method allowed to conduct cloning steps with successful products, which is actually noteworthy. In some cases, point mutations are difficult to perform and take several attempts even with standard cloning kits but here, due to the method chosen, the modification of 13 amino acids was achieved with no major obstacles.

However, the Three PCRs Method also had its drawbacks. It required many steps, and all of them dependent on the previous one. Moreover, it can be considered a time consuming procedure: the preparation and setting of the three PCRs programmed lasted roughly one hour and a half each, the two agarose gels and the extraction of the DNA from them took an average of an hour to perform, and any unsuccessful PCR1 or PCR2 forced the repetition of the process all over again.

Furthermore, as seen in Fig. 4.1 and Fig. 4.2, by initially cloning the huDHOase in several disjointed fragments of reduced size, unwanted amplified products derived from the reaction. The primers annealed in random complementary sites of the plasmid and started amplifying sequences that later were visible in agarose gels as thin and blurry bands of different sizes. This reduced the yield of the PCR reactions, as the primers were not exclusively dedicated to the amplification of the sought sequence.

Besides the loss in efficiency, cutting the gels for the purification of the DNA also resulted in a complicated and risky task. Many initial attempts for the 1st round were unsuccessful due to the obtaining of smeared gels like the one shown in Fig. 4.2b. When

cutting the gel, some DNA material that did not correspond to the PCR products desired was also purified and attached into the sequence, turning all the previous work fruitless.

Nevertheless, after the ligation step, the use of the QuikChange Site-directed Mutagenesis protocol for the second and third rounds resulted in an efficient intermediate choice of procedure in order to clone the huDHOase in less time while reducing the consumption of reagents and the number of dependent steps.

NaeI, key againsts time and expenditure burn up

Strikingly, the finding of NaeI as the only endonuclease that was able to cut in different patterns the mutated sequences, coming from each round of mutagenesis, was very fortunate. In this way, Serial Cloner's Virtual Cutter demonstrated its usefullness by filtering among the many restriction enzymes that exist in the market to discover the only one that fit the requirements, NaeI.

Despite the care taken when extracting the DNA from the gels, and the many controls performed during the transformation of TOP10 cells, the inclusion of a digestion step with the restriction enzyme NaeI resulted in the most appropriate approach to tackle any false positive. As a consequence, time and money were spared and the possible candidates were easily validated for their sending for Sanger sequencing.

Sequencing results of the huDHOase-ELF

When the first three rounds of mutagenesis were retrieved from Sanger sequencing, the analysis of the chromatograms showed the targeted bases successfully mutated just as expected, without any insertions or deletions that could cause the alterations of its sequence. Per contra, on the fourth round, besides the insertion of a histidine and the substitutions of the next three amino acids, another base was also substituted: g.4695C>T. This base, not belonging to the target sequence of the primers, was considered an undesirable mutation.

To determine if the base calling[33] for this nucleotide was correct, the chromatogram in Fig. 5.1 was inspected. Clearly, the curve representing the thymine in red has the correct shape and is exempt of noise [33]. However, to fully assure the legitimacy of this base read, the quality file also given by Stab Vida was checked.

Out of 60 in the Phred Score (Equation 5.1), this base position scored a 55, meaning that the probability of being mistakenly read is of 1 in 200,000 base calls [34]. Therefore,

[33] process done after sequencing in which every peak of the chromatogram is assigned one base

the silent mutation[34] was confirmed.

$$Q = -10 \log_{10} P \tag{5.1}$$

Where Q is the quality Phred Score and P is the base-calling error probability.

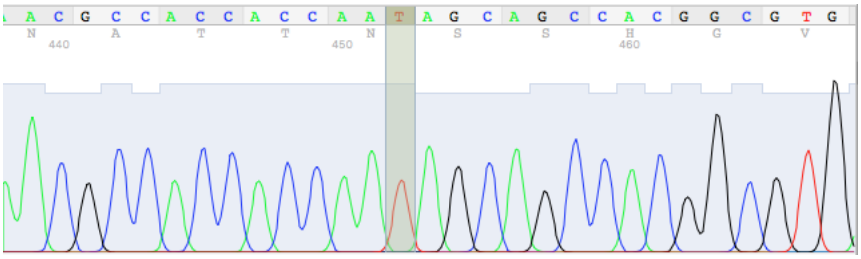


Fig. 5.1. Chromatogram showing the silent mutation of a cytosine into a thymine (g.4695C>T) in the huDHOase-ELF sequenced after the fourth round of mutagenesis.

However, thanks to codon degeneracy[35], the silent mutation and the two bases next to it coded for the same amino acid that was looked for during mutagenesis. Moreover, the abundance of asparagines that are coded by this generated triplet of bases in both human and *E.coli* species is 46% with respect to other asparagines translated by other codons [35], a very high abundance that allowed for the codon sequence to be left off intact. Therefore, no additional cloning step to reverse the mutation was performed and the silent mutation remained in the final huDHOase-ELF construct.

PEI as transfection vector for expression of mutant huDHOase-ELF

As shown in Fig. 4.10b, the expression of the GFP plasmid was only present in a small number of cells, meaning that the PEI ability of creating polyplexes with the DNA was poor. This can be due to two things:

1. The PEI used was branched, a type of PEI that has shown a remarkable transfection capacity in many studies in comparison with linear PEI [36] but presents high cytotoxic impact [37] which may be a possible explanation for the low number of cells expressing GFP.

[34] base substitution that generates a different codon but do not change its corresponding amino acid in the protein

[35] phenomenon by which several different codons code for the same amino acid in a protein, the redundancy of the genetic code

2. The ratio used for the induction of polyplexes was not properly chosen and the equilibrium of charges between the polymer (positive) and the plasmid (negative) was not achieved, hampering the cell uptake and endosomal escape[36] of the plasmid for its subsequent expression.

Despite these two possible limitations in the throughput, enough amount of protein was obtained for the activity assays. Furthermore, the ability of the cells to uptake the pOPIN-M-huDHOase-ELF plasmid and express it into a protein, presented preliminary indicia that the protein can be fairly expressed in mammalian cells even with such a major mutation on its sequence.

huDHOase-ELF and its singular purification

When IMAC was performed, and the histidine tagged huDHOase-ELF bound to MBP was eluted from the nickel cartridge, standard and recognizable peaks were obtained as shown in Fig. 4.11, with the addition of a conductivity peak between the ones representing the contaminants and the protein. This peak, characterized by its oscillatory behaviour, is an artifact possibly caused by a bubble of air passing through the UV detector.

Aside from this punctual artifact, the HPLC was used again for the separation of the huDHOase-ELF from the 6xHis Tag + MBP and, once again, ordinary and standard peaks were retrieved and taken as indicators of the proper sorting of both elements. However, when the gel with the huDHOase-ELF samples after digestion was checked, a second band below the expected one (in between the 40 and 50 kDa markers) appeared in wells 2, 5, 7, 9 and 11. There are two hypotheses to explain this unexpected bands in the gel:

1. The SDS-PAGE gel might not have been properly polymerized and the double band was caused by the discontinuous matrix of the polymer. Nevertheless, this hypothesis is partially discarded as the rest of the bands shown in the gel do not present a duplication like the one observed in those wells.

2. The bands below may indicate the presence of protein with lower amount of amino acids and therefore, with lower molecular weight. This may be due to the possible degradation of the C-terminal extension of the huDHOase-ELF by the proteases used during the purification process. This terminal extension, as described by Grande-García *et al.*, is a fundamental part for the protein as it is required for its solubility and it is very sensitive to cleavage [8]. Therefore, it is possible that by using a non-optimal amount of proteases during its purification, some of the protein might not have its amino acid sequence intact.

[36] pH-triggered process by which the polyplexes rupture the membrane of the endosome and are released into the cytoplasm of the cell

After this second band was visualized, the purification was continued with special care for the last size exclusion chromatography step. In this case, it was actually the gel the one showing ordinary results with unique bands in each well between the 40 and 50 kDa size markers. When the chromatogram retrieved from the HPLC system was examined, the peak corresponding to the elution of the protein remained within a range between 35 and 67 kDa, not the expected size of the huDHOase in its dimer conformation. This elution profile, however, fits better the size of only one subunit of the huDHOase (approx. 42.8 kDa) and implies the behaviour of the protein as a monomer in solution while the purification took place, an interesting premise taken into account later, in addition to the hypothesis of the degradation of the C-terminal, when evaluation of the catalytic rate of the protein was performed.

Moreover, despite the protein being successfully isolated from other components and properly purified, the yield of this procedure was very low. After concentration of the protein, only 60 μl were retrieved. A volume that forced to choose between assaying the protein's activity or doing crystallization screening to later solve its structure by X-ray diffraction [8]. Consequently, special care should be taken during future purifications of the protein, as the loss of protein bound to contaminants, and after concentration, confirms its delicate and unstable nature during purification.

Activity assays, considerations on methodology and kinetic results

Before a comparison between the wild type huDHOase and the chimera huDHOase-ELF can be drawn, the activity coming from the wt-huDHOase assays was compared with the previous work done about this protein. In the study of huDHOase's kinetic parameters done by Grande-García *et al.*, both catalytic constants, k_{cat} = 191.4 ± 2.6 min^{-1} for the forward reaction and k_{cat} = 349.8 ± 10.4 min^{-1} for the reverse direction [8], were higher than the ones obtained in this project (113.5 ± 0.5 min^{-1} and 181.8 ± 13.2 min^{-1} respectively). However, just like in their study, the k_{cat} for the forward reaction presents a lower value in comparison with the one corresponding to the reverse direction.

There are two possible explanations for the difference between the huDHOase's k_{cat} obtained in Grande-García's *et al* study and the one here reported.

On the one hand, the wt-huDHOase used in this experiment was frozen and then thawed repeatedly, as it was used for many past experiments before this project. As a consequence, the temperature changes suffered by the protein might have been detrimental and could explain the reduction of its activity [38].

On the other hand, another potential reason is the pH at which the activity assays were

performed, as huDHOase activity is pH dependent [8]. While Grande-García *et al.* used a pH of 5.5 and 8 for the forward and reverse direction respectively, in this project a pH of 5 and 8.5 were employed, as both the wild type and mutant proteins needed to be tested at the same pH and going into the ones used for maximum catalytic activity could have a deleterious effect on the mutant.

While other colorimetric and radioactive methods employed to detect the formation of CA-asp [39], or both directions of the reaction [40], are considered more sensitive and appropriate for activity assays, other works have reported how all of them showed similar results in the determination of huDHOase k_{cat} [41]. Due to this fact, the triplicate detection of DHO formation/consumption in each assay was chosen and with that, many sources of error or strenuous steps were avoided. However, to guarantee the validity of the results obtained with this implemented method, there are two aspects that were taken into account: (1) in addition to the pipetting error (2%), the spectrophotometer used has 1% of random error in its absorbance detection; (2) the linear response of the absorbance of DHO remains between the 0-54 mM range. To counteract both, assays were done in triplicates and the concentration of protein and substrate added to the reaction were adjusted so no saturation of the detector could occur.

However, as many elements in the sample, apart from tryptophan, present absorbance at 230 nm (Tris Buffer, DHOase, EDTA, etc.), it must be taken into account that the production/consumption of DHO was not the exclusive measure at each time point and that *chemical noise* might have affected the data [42].

The implications behind huDHOase-ELF's activity results

As seen in Fig. 4.18a and Fig. 4.19a, the spectrophotometer provided the expected measures of absorbance when the wild type huDHOase produced or consumed DHO respectively. In the case of the readings for the mutant huDHOase-ELF, with the exception of assay 3 in the forward reaction (Fig. 4.18b), the fluctuations in the measures were minimal and the slope of the graphs from both directions manifest the, however small, activity of the protein.

Indeed, as presented in Fig. 4.17 and Fig. 4.20, parted y axis were needed to represent the negligible activity of the huDHOase-ELF next to the wt-huDHOase. With a 67-fold difference in the forward direction and 31 in the reverse direction, the chimera developed does not hold the same catalytic activity exhibited by the wild type. Nonetheless, the slopes calculated in all six essays performed on the mutant all gave different values from zero, demonstrating that although small, the ability to transform the substrate CA-asp into the product (Fig. 4.18b) and vice versa (Fig. 4.19b) remained. The main reason accountable for this dramatic decrease in the activity is, as expected, the considerable mutation

performed to the sequence and its effects on the movement mechanism of the flexible loop. However, before elaborating into this assumption, two important consideration must be taken into account:

- As discussed previously, the C-terminal extension of the protein might have suffered a cleavage during purification, which could be translated into structural and property changes that affected its catalytic activity besides the interchange of flexible loop with ecDHOase conducted.
- After size exclusion chromatography, the protein eluted in fractions corresponding to a given size range but its exact molar mass was not determined as for that, MALS[37] (Multi-Angle Light Scattering) should be coupled to the HPLC system after size exclusion chromatography. Therefore, as the protein was not confirmed to be a dimer, it is possible that the first amino acid mutations in the flexible loop not only affected the mechanism of catalysis of the enzyme but its ability to oligomerize[38]. And, as the estimated loss of activity when huDHOase is in monomer conformation (according to work pending for publication under the name of Ramón-Maiques) is half the activity the dimer presents, it should be taken into account when examining the kinetic data.

Despite the fact these two preliminary remarks could not be examined further with the data obtained during the project, the catalytic activity of the chimeric huDHOase-ELF was so tremendously affected that both considerations are not enough to justify it. Therefore, the mutation of the human flexible loop into the *E.coli* sequence can be described as the major responsible for this decrease in activity.

As seen in Fig. 4.17, with a 67-fold decrease in activity for the forward reaction, the chimeric huDHOase-ELF demonstrates that the flexible loop in humans is irreplaceable. Despite having some similarities with *E.coli*'s catalytic mechanism, both huDHOase-ELF k_{cat} here reported prove that the change in amino acid sequence can alter the mechanical and catalytic behavior of the loop.

An important detail regarding this sequence identity involves the residue Thr109 in ecDHOase and its equivalent in huDHOase, Thr1562. This threonine, positioned at the tip of the loop, is reported to be a key interacting element with inhibitor molecules and the substrate, forming a hydrogen bond with CA-asp in *E.coli* [13] and humans [8]. This hydrogen bond would allow the proper confinement of the substrate into the active site favoring and providing stability to the catalytic reaction while avoiding unwanted interactions [12]. As this function seems important for the catalysis, many species have this threonine conserved (*E.coli, Homo Sapiens, P.gingivalis* etc.) but, as demonstrated with

[37] technique employed for the precise measurement of molar mass and characterization of oligomers

[38] process in which a protein assembles into oligomers by the polymerization of its monomer constituents

the activity of huDHOase-ELF obtained, even if the residue is left intact, insertions and mutations may disorder the loop, inducing steric clashes that could hinder the reactions between the atoms in the active site and this threonine to take place.

Several other relationships between the flexible loop and the elements of the protein can also be speculated to end up affected. As explained on Section 1.1, if the mutation of the flexible loop or the residues in charge of the dimer interface caused the change of the movement mechanism, then the communication between subunits might have been hampered and the cooperativity[39] between subunits impaired [12]. On the other side, if the bridging ligands between the metal cations were displaced, the whole arrangement of the active site might have been altered, with the consequences that would entails.

Nonetheless, both hypotheses on the lack of stabilization of the flexible loop into its **IN** conformation while sequestering the substrate/inhibitor, and the alteration of its spatial configuration, need to be examined once the 3D structure of the protein is solved as for now the data obtained only provided with the possibility of speculation.

Finally, with the catalytic activity reported for this mutant, it is somehow demonstrated that, even with similar sequences, the flexible loop from one DHOase species cannot be substituted with another loop from other species. Other elements, apart from the loop's residues, are actively participating into the catalysis and their interactions with the loop are crucial for its mechanism.

[39] change of shape and affinity of a protein subunit when a ligand binds to the active site of other of its subunits

6. CONCLUSIONS AND FUTURE WORK

With the experiments performed during the work here presented, the design of a feasible and robust cloning method for the mutation of huDHOase flexible loop into a new sequence of 13 amino acids coming from *E.coli* was achieved. By using a combination of The Three PCRs Method and QuikChange Site-directed Mutagenesis, it was possible to create a chimera between both species with additional features in its structure for its efficient purification during the implementation of chromatography techniques as IMAC and SEC [15].

Nonetheless, experiments must continue, as the huDHOase-ELF obtained by expression of the plasmid on mammalian cells presented an unstable nature during purification and its characteristics as monomer or dimer, as well as the integrity of its complete sequence, were unable to be described with the experiments performed and the data retrieved. Due to this, expression using induction instead of transfection, and the use of IMAC followed by coupling MALS to the size exclusion chromatography technique, must be performed. These two different approaches in the expression and purification strategy might help to answer if the reduced amount of protein obtained was caused by the inefficiency of PEI as transfection vector, or if its unstable behaviour and its monomer-like elution profile during purification was caused by an alteration of its structure. Therefore, the molar mass of the protein must be first described by MALS and, after that, the mutations conducted must be inspected in case they affected to the oligomerization of the protein.

But, most importantly, the overall behaviour of the protein is clearly altered and the experiments here described point at the loop as the main suspect responsible the changes observed on its solubility and oligomer formation. Interestingly, the negligible activity shown by the chimera huDHOase-ELF availed the statements pointing at the flexible loop as a crucial element for solubility and oligomer stability of the protein and, at the same time, as the main structure behind the huDHOase catalysis. This flexible loop might be so exclusive for each species that, even with similar movement and catalytic mechanism, the sequence of amino acids is so refined for each species that the studies previously done on *E.coli* DHOase do not provide with enough knowledge that could be extrapolated to human. In the case, for example, of the Plasmodium parasite responsible for malaria, this would imply the possibility of designing species-specific inhibitors that affect only the DHOase of interest, as DHOases might share similar catalytic behaviour but with major differences in its mechanical/biochemical mechanism.

Moved by the referred specificity of the flexible loop mechanism, crystallization screening is planned to be performed in order to solve the conformation of the chimera huDHOase-ELF by X-ray crystallography in the near future [43]. With the three-dimensional structure given by this technique, interesting data could be extracted from the arrangement of its binding pocket [15] and models of the loop's movement could explain more specifically its interactions with other elements in the protein.

Furthermore, once the behaviour of this chimera is incorporated to the previous knowledge compiled from other mutants, some clues on huDHOase crucial elements could be derived aiding in the search for robust cytotoxic inhibitors. These inhibitors, acting as cytotoxic agents targeting DHOases by resilient interactions and halting its catalytic mechanism, could become potential therapeutic tools in the future battle against cancer cells or parasites as Plasmodium.

Finally, as this work will be presented shortly as part of the 26th International Conference on Arginine and Pyrimidines and the 1st International Conference on Amino Acids and Nucleotides (ICAP | ICAN2018) at ShanghaiTech University, it is expected that the discussion over the project will enlighten the results obtained and new approaches in the methodology might be suggested.

7. SOCIO-ECONOMIC IMPACT

7.1. Analysis of the project's envisioned repercussions

In the case of the work here presented, with the crystallographic analysis of the synthesized huDHOase-ELF that will be continued after the catalytic results obtained, three main spheres might suffer an impact in a domino effect manner:

Academia, field impact.

The methodology followed for the obtaining of the mutant huDHOase, as well as the activity results preceding the crystallographic work, represent key elements in the knowledge contribution this work is providing to the scientific community. The better understanding of pyrimidine biosynthesis might encourage others to research on other CAD domains or the metabolic alteration cells suffer in this metabolic pathway when turned neoplastic.

The impact on the economy of drug development and drug efficiency.

With the knowledge acquired, and the confidence of the results, pharmacological companies might find the prospect of designing an inhibitor against huDHOase attractive. However, how different can the huDHOase-ELF make the drug market? By the throughout and careful study of the target, better synthetic compounds can be designed. And this, as basic as it may appear, makes a real difference. Current studies show how 90% of antitumoral drugs, even in combination with other therapies like chemotherapy and radiotherapy, fail to pass the so-called Valley of Death [44], the stage in which most of the drugs are unapproved due to its effectiveness and unsatisfactory results. And, to bring those drugs to the edge of the valley even before they fall, pharma companies spend 1.55 billion of dollars [45]. Due to the risk that drug development entails, and the high monetary losses pharma companies face, only by providing them with a feasible and credible venture proposal, investments on experimental drugs that cross the Valley could achieve success.

Furthermore, by a sensible combination between this potential drug and other anticancer treatments, a proper use of healthcare resources would be executed, decreasing the costs public healthcare system abide on useless and hopeless treatments.

Patients, the human impact

And finally, and most importantly, this study could ultimately have an impact on human lives. A more careful study on the target would produce a more efficient inhibitor that, in a particular combination with other treatments, would provide cancer patients with a personalized approach in the battle against the disease. This, in fact, means an intangible impact on the patients regarding two aspects:

- *The lessening on their distress and shortening of the suffering due to current strenuous treatments.* With the discovery of better targets, the misuse of cytotoxic treatments like chemotherapy and radiotherapy would be overcame, and the pain caused by the adverse effects related to them could be avoided.
- *Personalized medicine and increased expectations on the treatment*. By analyzing the patient prospect and their needs, the combination of traditional treatments with enzyme inhibitors like PALA or inhibitors designed against huDHOase, could aid in the implementation of personalized medicine and the confidence on a good result would help in the psychological battle against the disease. This, as pointless as it might sound, can also make a difference in the way patients confront their state.

7.2. Estimated budget of the project

The estimated budget for the accomplishment of the project presented can be divided in three sections. First, the approximate expenditure on common fungible consumables and reagents followed by the working hours carried out by both the student and supervisor tutor, and lastly a detailed list on the indispensable elements needed to replicate the cloning, expression and purification of the protein.

An aspect to take into account is the use of free software and student licenses for all the programs used in this project: Serial Cloner (Freeware), 4peaks (Freeware), PyMOL (academic license); except for Prism. GraphPad Prism was provided by the CBMSO laboratory where the project was done and which cost (170 €) is included in Table 7.1.

Description	Estimated net expenditure per month (€)	Number of Months	Total Cost (€)
Fungible goods (pipettes, tubes, eppendorfs, buffers etc.) software and equipment of the laboratory	700	5	3500

TABLE 7.1. ESTIMATED NET COST OF FUNGIBLE GOODS AND EQUIPMENT DURING THE MONTHS THE PROJECT LASTED

Category	Hours	Gross Cost Per Hour (€)	Total Cost (€)
Biomedical engineer	620	10	8680
Research Investigator	70	15	1050

TABLE 7.2. ESTIMATED GROSS COST OF WORKING HOURS

Description	Units	Cost per unit (€)	Total cost (€)
pOPINM-M-huDHOase (Addgene)	1	94.5	95.4
NEB Buffer 10X (New England Biolabs)	1	18.8	18.8
BSA (New England Biolabs)	1	24.9	24.9
HindIII (New England Biolabs) 10,000 units	1	31.8	31.8
KpnI (New England Biolabs) 4,000 units	1	36.1	36.1
Primers (Sigma)	10	5.94	59.4
dNTPs (Thermo Fisher) 4x 250 µl	1	254	254
Phusion 2000 units/ml Buffer Phusion 5X DMSO (New England Biolabs)	1	91.8	91.8
Agarose (CONDA)	1	42	42
GelRed (Biotium)	1	85.75	85.75
QIAquick Gel Extraction Kit 50 rxns (Qiagen)	1	101.2	101.2
GeneJET Plasmid Miniprep Kit (Thermo Fisher)	1	62.75	62.75
In-Fusion HD (Clontech)	1	198.17	198.17
Expand Long System Polymerase Buffer Expand Long 10X (Sigma)	1	145	145
QuikChange Kit (Agilent)	1	355	355
One Shot TOP10 E.coli competent cells (Thermo Fisher)	1	548	548
NaeI (New England Biolabs)	1	33.55	33.55
Sanger sequencing (Stab Vida)	4	3.5	14
HEK293S-GnTI (ATCC)	1	590	590
Freestyle medium (Gibco)	1	166	166
Polyethylenimine (Sigma)	1	86.5	86.5
UltraDOMA medium (Lonza)	1	26.25	26.25
pCMV-GFP Addgene	1	55.75	55.75
Nickel Cartridge (ABT) AF6Ni-Ctg5-1	1	97.8	97.8
PreScission protease (GE Healthcare) 500 u	1	278.76	278.76
Amicon Ultra filters system (Sigma) 5kDa	1	49.5	49.5
Amicon Ultra filters system (Sigma)10 kDa	1	107	107
Superdex 200 10/300 column was used (GE Healthcare)	1	1474	1474
Acrylamide (Sigma) Powder 100 g	1	58	58
SDS 10% H2O (Sigma)	1	39.75	39.75

TABLE 7.3. DETAILED LIST OF INDISPENSABLE REAGENTS USED AND THEIR NET COST

Description	Total Cost (€)
Fungible consumables and equipment of the laboratory	3500
Working Hours	9730
Indispensable Reagents	5226.93
IVA (21%)	1832.65
TOTAL	20289.58

TABLE 7.4. FINAL ESTIMATED BUDGET OF THE PROJECT

BIBLIOGRAPHY

[1] D. L. Nelson and M. M. Cox, *Principles of Biochemistry*, 9. 2013, vol. 53, pp. 1689–1699, ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004.

[2] Z. Martins *et al.*, "Extraterrestrial nucleobases in the Murchison meteorite", *Earth and Planetary Science Letters*, vol. 270, no. 1-2, pp. 130–136, 2008, ISSN: 0012821X. DOI: 10.1016/j.epsl.2008.03.026.

[3] J. Koch *et al.*, "CAD mutations and uridine-responsive epileptic encephalopathy", *Brain*, vol. 140, no. 2, pp. 279–286, 2017, ISSN: 14602156. DOI: 10.1093/brain/aww300.

[4] F. D. Sigoillot, J. A. Berkowski, S. M. Sigoillot, D. H. Kotsis and H. I. Guy, "Cell cycle-dependent regulation of pyrimidine biosynthesis", *Journal of Biological Chemistry*, vol. 278, no. 5, pp. 3403–3409, 2003, ISSN: 00219258. DOI: 10.1074/jbc.M211078200.

[5] D. R. Evans and H. I. Guy, "Mammalian pyrimidine biosynthesis: Fresh insights into an ancient pathway", *Journal of Biological Chemistry*, vol. 279, no. 32, pp. 33 035–33 038, 2004, ISSN: 00219258. DOI: 10.1074/jbc.R400007200.

[6] M. Moreno-Morcillo *et al.*, "Structural Insight into the Core of CAD, the Multifunctional Protein Leading De Novo Pyrimidine Biosynthesis", *Structure*, vol. 25, no. 6, 912–923.e5, 2017, ISSN: 18784186. DOI: 10.1016/j.str.2017.04.012.

[7] G. R. Stark and P. A. Bartlett, "Design and use of potent, specific enzyme inhibitors", *Pharmacol Ther*, vol. 23, no. 1, pp. 45–78, 1983. DOI: 0163-7258(83)90026-8[pii]. [Online]. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve%7B%5C&%7Ddb=PubMed%7B%5C&%7Ddopt=Citation%7B%5C&%7Dlist%7B%5C_%7Duids=6361807.

[8] A. Grande-García, N. Lallous, C. Díaz-Tejada and S. Ramón-Maiques, "Structure, functional characterization, and evolution of the dihydroorotase domain of human CAD", *Structure*, vol. 22, no. 2, pp. 185–198, 2014, ISSN: 09692126. DOI: 10.1016/j.str.2013.10.016.

[9] B. G. Ng *et al.*, "Biallelic mutations in CAD, impair de novo pyrimidine biosynthesis and decrease glycosylation precursors", *Human Molecular Genetics*, vol. 24, no. 11, pp. 3050–3057, 2014, ISSN: 14602083. DOI: 10.1093/hmg/ddv057.

[10] J. B. Thoden, G. N. Phillips, T. M. Neal, F. M. Raushel and H. M. Holden, "Molecular structure of dihydroorotase: a paradigm for catalysis through the use of a binuclear metal center.", *Biochemistry*, vol. 40, no. 24, pp. 6989–6997, 2001, ISSN: 0006-2960. DOI: 10.1021/bi010682i.

[11] M. Lee, M. J. Maher, R. I. Christopherson and J. M. Guss, "Kinetic and structural analysis of mutant Escherichia coli dihydroorotases: A flexible loop stabilizes the transition state", *Biochemistry*, vol. 46, no. 37, pp. 10 538–10 550, 2007. DOI: Doi10.1021/Bi701098e.

[12] M. Lee, C. W. Chan, J. M. Guss, R. I. Christopherson and M. J. Maher, "Dihydroorotase from Escherichia coli: Loop movement and cooperativity between subunits", *Journal of Molecular Biology*, vol. 348, no. 3, pp. 523–533, 2005, ISSN: 00222836. DOI: 10.1016/j.jmb.2005.01.067.

[13] M. Lee *et al.*, "Structures of Ligand-free and Inhibitor Complexes of Dihydroorotase from Escherichia coli: Implications for Loop Movement in Inhibitor Design", *Journal of Molecular Biology*, vol. 370, no. 5, pp. 812–825, 2007, ISSN: 00222836. DOI: 10.1016/j.jmb.2007.05.019.

[14] R. I. Christopherson and M. E. Jones, "Interconversion of carbamayl-L-aspartate and L-dihydroorotate by dihydroorotase from mouse Ehrlich ascites carcinoma.", *The Journal of biological chemistry*, vol. 254, no. 24, pp. 12 506–12 512, 1979, ISSN: 0021-9258. [Online]. Available: http://www.jbc.org/content/254/24/12506.short%7B%5C%%7D5Cnpapers2://publication/uuid/B8B08B95-48E4-4EFB-90B9-D4DB68657E31.

[15] A. Ruiz-Ramos, N. Lallous, A. Grande-García and S. Ramón-Maiques, "Expression, purification, crystallization and preliminary X-ray diffraction analysis of the aspartate transcarbamoylase domain of human CAD", *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, vol. 69, no. 12, pp. 1425–1430, 2013, ISSN: 17443091. DOI: 10.1107/S1744309113031114.

[16] D. Roy, G. J. Steyer, M. Gargesha, M. E. Stone and L. Wilson, "Design, synthesis and bioactivity of novel inhibitors of E. coli aspartate transcarbamoylase", in *Bioorganic & medicinal chemistry letters*, 7, vol. 17, 2007, pp. 2086–2090, ISBN: 0130149616. DOI: 10.1016/j.bmcl.2006.12.050.

[17] J. L. Grem, S. A. King, P. J. O'Dwyer and B. Leyland-Jones, "Biochemistry and clinical activity of N-(phosphonacetyl)-L-aspartate: a review.", *Cancer research*, vol. 48, no. 16, pp. 4441–4454, 1988, ISSN: 00085472. [Online]. Available: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed%7B%5C&%7Did=3293772%7B%5C&%7Dretmode=ref%7B%5C&%7Dcmd=prlinks%7B%5C%%7D5Cnpapers2://publication/uuid/C2308983-E9D9-46EE-95BB-4BED1ABD9577.

[18] K. K. Tsuboi, H. N. Edmunds and L. K. Kwong, "Selective Inhibition of Pyrimidine Biosynthesis Proliferative Growth of Colonie Cancer Cells1", *Cancer Research*, vol. 37, no. September, pp. 3080–3087, 1977.

[19] M. Lee *et al.*, "Structures of Ligand-free and Inhibitor Complexes of Dihydroorotase from Escherichia coli: Implications for Loop Movement in Inhibitor Design", *Journal of Molecular Biology*, vol. 370, no. 5, pp. 812–825, 2007, ISSN: 00222836. DOI: 10.1016/j.jmb.2007.05.019.

[20] Z. M. Gomez and P. K. Rathod, "Antimalarial activity of a combination of 5-fluoroorotate and uridine in mice", *Antimicrobial Agents and Chemotherapy*, vol. 34, no. 7, pp. 1371–1375, 1990, ISSN: 00664804. DOI: 10.1128/AAC.34.7.1371.

[21] R. I. Christopherson *et al.*, "Mercaptan and dicarboxylate inhibitors of hamster dihydroorotase", *Biochemistry*, vol. 28, no. 2, pp. 463–470, 1989, PMID: 2565732. DOI: 10.1021/bi00428a009. [Online]. Available: https://doi.org/10.1021/bi00428a009.

[22] J. Brooke *et al.*, "Cytotoxic Effects of Dihydroorotase Inhibitors upon Human CCRF-CEM Leukemia", *Cancer Research*, vol. 50, no. 24, pp. 7793–7798, 1990, ISSN: 15387445.

[23] V. Revadigar *et al.*, "Enzyme Inhibitors Involved in the Treatment of Alzheimer's Disease", in *Drug Design and Discovery in Alzheimer's Disease*, Elsevier, 2014, pp. 142–198, ISBN: 9780128039595. DOI: 10.1016/B978-0-12-803959-5.50003-9. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/B9780128039595500039.

[24] V. Chopra *et al.*, "The sirtuin 2 inhibitor AK-7 is neuroprotective in Huntington's disease mouse models.", *Cell reports*, vol. 2, no. 6, pp. 1492–7, Dec. 2012, ISSN: 2211-1247. DOI: 10.1016/j.celrep.2012.11.001. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23200855%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3534897.

[25] H. K. Weir *et al.*, "Heart Disease and Cancer Deaths — Trends and Projections in the United States, 1969–2020", *Preventing Chronic Disease*, vol. 13, p. 160 211, 2016, ISSN: 1545-1151. DOI: 10.5888/pcd13.160211. [Online]. Available: http://www.cdc.gov/pcd/issues/2016/16%7B%5C_%7D0211.htm.

[26] N. Hamajima, T. Saito, K. Matsuo and K. Tajima, "Competitive amplification and unspecific amplification in polymerase chain reaction with confronting two-pair primers", *Journal of Molecular Diagnostics*, vol. 4, no. 2, pp. 103–107, 2002, ISSN: 15251578. DOI: 10.1016/S1525-1578(10)60688-5.

[27] P. Y. Lee, J. Costumbrado, C.-Y. Hsu and Y. H. Kim, "Agarose Gel Electrophoresis for the Separation of DNA Fragments", *Journal of Visualized Experiments*, no. 62, pp. 1–5, 2012, ISSN: 1940-087X. DOI: 10.3791/3923. [Online]. Available: http://www.jove.com/video/3923/.

[28] P. a. Longo, J. M. Kavran, M.-s. Kim and D. J. Leahy, "Transient Mammalian Cell Transfection wtih Polyethylenimine", *Methods Enzymology*, vol. 529, pp. 227–240, 2014. DOI: 10.1016/B978-0-12-418687-3.00018-5.Transient.

[29] K. Weber and M. Osborn, "The Reliability of Molecular Weight Determinations Sulf ate-Polyacrylamide Gel Electrophoresis", *The Journal of Biological Chemistry*, vol. 244, no. 16, pp. 4406–4412, 1969, ISSN: 0021-9258, 1083-351X. [Online]. Available: http://www.jbc.org/content/244/16/4406.short.

[30] M. M. Bradford, "A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of dye-binding", *Anal. Biochem.*, vol. 72, pp. 248–254, 1976.

[31] E. G. Sander, L. D. Wright and D. B. McCormick, "Evidence for function of a metal ion in the activity of dihydroorotase from Zymobacterium oroticum.", *The Journal of biological chemistry*, vol. 240, no. 9, pp. 3628–3630, 1965, ISSN: 00219258.

[32] D. F. Swinehart, "The Beer-Lambert Law", *Journal of Chemical Education*, vol. 39, no. 7, p. 333, 1962, ISSN: 0021-9584. DOI: 10.1021/ed039p333. [Online]. Available: http://pubs.acs.org/doi/abs/10.1021/ed039p333.

[33] P. Cliften, "Chapter 7 - base calling, read mapping, and coverage analysis", in *Clinical Genomics*, S. Kulkarni and J. Pfeifer, Eds., Boston: Academic Press, 2015, pp. 91–107, ISBN: 978-0-12-404748-8. DOI: https://doi.org/10.1016/B978-0-12-404748-8.00007-1. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780124047488000071.

[34] C. Ledergerber and C. Dessimoz, "Base-calling for next-generation sequencing platforms", *Briefings in Bioinformatics*, vol. 12, no. 5, pp. 489–497, 2011, ISSN: 14675463. DOI: 10.1093/bib/bbq077.

[35] J. Athey *et al.*, "A new and updated resource for codon usage tables", *BMC Bioinformatics*, vol. 18, no. 1, pp. 1–10, 2017, ISSN: 14712105. DOI: 10.1186/s12859-017-1793-7.

[36] A. Bragonzi *et al.*, "Comparison between cationic polymers and lipids in mediating systemic gene delivery to the lungs", *Gene Therapy*, vol. 6, no. 12, pp. 1995–2004, 1999, ISSN: 09697128. DOI: 10.1038/sj.gt.3301039.

[37] K. V and O. Y, "Cytotoxic Impacts of Linear and Branched Polyethylenimine Nanostructures in A431 Cells", *BioImpacts*, vol. 1, no. 1, pp. 23–30, Jan. 2017. DOI: 10.5681/BI.2011.004. [Online]. Available: http://bi.tbzmed.ac.ir/Abstract/BI%7B%5C_%7D20120818132636.

[38] R. M. Daniel and M. J. Danson, "Temperature and the catalytic activity of enzymes: A fresh understanding", *FEBS Letters*, vol. 587, no. 17, pp. 2738–2743, 2013, ISSN: 00145793. DOI: 10.1016/j.febslet.2013.06.027. [Online]. Available: http://dx.doi.org/10.1016/j.febslet.2013.06.027.

[39] L. M. Prescott and M. E. Jones, "Modified methods for the determination of carbamyl aspartate", *Analytical Biochemistry*, vol. 32, no. 3, pp. 408–419, 1969, ISSN: 10960309. DOI: 10.1016/S0003-2697(69)80008-4.

[40] R. Christopherson, T. Matsuura and M. Jones, "Radioassay of dihydroorotase utilizing ion-exchange chromatography", *Analytical biochemistry*, vol. 89, no. 1, pp. 225–234, Aug. 1978, ISSN: 0003-2697. DOI: 10.1016/0003-2697(78)90745-5. [Online]. Available: https://doi.org/10.1016/0003-2697(78)90745-5.

[41] R. Alba, G. Araceli and R. Santiago, "Dihydroorotase domain of human cad", in *Encyclopedia of Inorganic and Bioinorganic Chemistry*. American Cancer Society, 2015, pp. 1–16, ISBN: 9781119951438. DOI: 10.1002/9781119951438.eibc2321. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119951438.eibc2321.

[42] P. F. Liu, L. V. Avramova and C. Park, "Revisiting absorbance at 230 nm as a protein unfolding probe", *Analytical Biochemistry*, vol. 389, no. 2, pp. 165–170, 2009, ISSN: 00032697. DOI: 10.1016/j.ab.2009.03.028. [Online]. Available: http://dx.doi.org/10.1016/j.ab.2009.03.028.

[43] E. Ennifar, "X-ray Crystallography as a Tool for Mechanism-of-Action Studies and Drug Discovery", *Current Pharmaceutical Biotechnology*, vol. 14, no. 5, pp. 537–550, Nov. 2013, ISSN: 13892010. DOI: 10.2174/13892010140513111110482 4. [Online]. Available: http://www.eurekaselect.com/openurl/content.php?genre=article%7B%5C&%7Dissn=1389-2010%7B%5C&%7Dvolume=14%7B%5C&%7Dissue=5%7B%5C&%7Dspage=537.

[44] D. J. Adams, "The valley of death in anticancer drug development: A re-assessment", *Trends in Pharmacological Science*, vol. 33, no. 4, pp. 173–180, 2012. DOI: 10.1016/j.tips.2012.02.001.The.

[45] M. Hay, D. W. Thomas, J. L. Craighead, C. Economides and J. Rosenthal, "Clinical development success rates for investigational drugs", *Nature Biotechnology*, vol. 32, p. 40, Jan. 2014. [Online]. Available: http://dx.doi.org/10.1038/nbt.2786%20http://10.0.4.14/nbt.2786%20https://www.nature.com/articles/nbt.2786%7B%5C#%7Dsupplementary-information.